Trends In Amplification

VOLUME 6, NUMBER 3, 2002

Methods and Applications of the Audibility Index in Hearing Aid Selection and Fitting

Amyn M. Amlani, MS,* Jerry L. Punch, PhD,* and Teresa Y. C. Ching, PhD†

During the first half of the 20th century, communications engineers at Bell Telephone Laboratories developed the articulation model for predicting speech intelligibility transmitted through different telecommunication devices under varying electroacoustic conditions. The profession of audiology adopted this model and its quantitative aspects, known as the Articulation Index and Speech Intelligibility Index, and applied these indices to the prediction of unaided and aided speech intelligibility in hearing-impaired listeners. Over time, the calculation methods of these indices—referred to collectively in this paper as the Audibility Index—have been continually refined and simplified for clinical use. This article provides (1) an overview of the basic principles and the calculation methods of the Audibility Index, the Speech Transmission Index and related indices, as well as the Speech Recognition Sensitivity Model, (2) a review of the literature on using the Audibility Index to predict speech intelligibility of hearing-impaired listeners, (3) a review of the literature on the applicability of the Audibility Index to the selection and fitting of hearing aids, and (4) a discussion of future scientific needs and clinical applications of the Audibility Index.

"Articulation measurement is a tedious and expensive business, and considerable attention has been devoted to the development of a computational device to replace, or at least to supplement, the laborious testing procedures"—LICKLIDER AND MILLER (1951, p. 1055)

Introduction

The model of articulation theory was developed at Bell Telephone Laboratories as a means to predict speech signals transmitted through different telecommunication devices under varying electroacoustic conditions (French and Steinberg, 1947). Specifically, this model assumes that the intelligibility of speech through any communication system can be described using weighted measurements of the speech-frequency regions audible to the listener. In the model, audible speech

From the *Department of Audiology and Speech Sciences, Michigan State University, East Lansing, Michigan; and the †National Acoustic Laboratories, 126 Greville Street, Chatswood, NSW 2067, Australia

Correspondence: Amyn M. Amlani, Department of Audiology and Speech Sciences, Michigan State University, East Lansing, MI 48824 E-mail: amlaniam@msu.edu

©2002 Westminster Publications, Inc., 708 Glen Cove Avenue, Glen Head, NY 11545, U.S.A.

cues in a given frequency band determine the amount of information in the band, and thus its contribution to the overall intelligibility of speech.

As a way to quantify the relationship between the audible speech cues and intelligibility, the Articulation Index was developed as an intermediate step. The Articulation Index is defined as "a weighted fraction representing, for a given speech channel and noise condition, the effective proportion of the normal speech signal which is available to a listener for conveying speech intelligibility" (Kryter, 1962[a], p. 1689). This calculated proportion ranges from 0.0 to 1.0 and is used subsequently to predict the speech intelligibility.

The Articulation Index has had, and continues to have, major practical implications for the practice of audiology. Clinical use of the Articulation Index allows audiologists to predict speech intelligibility under unaided listening conditions, as well as the benefit to be derived from a hearing aid, or aids, by comparing predicted intelligibility performance under unaided and aided conditions (eg, Rankovic, 1991; Studebaker and Sherbecoe, 1993). Many audiologists use count-the-dot-audiograms (Mueller and Killion, 1990; Humes, 1991; Pavlovic, 1991), a simplification of the Articulation Index, to demonstrate to their patients the effects of hearing impairment on speech understanding without and with amplification.

The Articulation Index has also played an important historical role in procedures designed to prescribe the frequency-gain characteristics of linear hearing aids (eg, Humes, 1986; Rankovic, 1991). The basic goal of these threshold-based methods is to amplify the speech spectrum so that the long-term average speech in each band is 15 to 18 dB above threshold at each frequency. According to articulation theory, this allows the full 30-dB range of speech to be audible (ie, detectable) at each frequency and provides maximum speech intelligibility (Humes, 1986; Rankovic, 1991; Humes and Halling, 1994). The Articulation Index model was adopted by the National Acoustic Laboratories in Australia in the derivation of one of the most recent methods of fitting nonlinear hearing aids (Dillon, 1999; Byrne *et al.*, 2001).

Articulation Index-based methods are incorporated into various commercially available hearing aid analyzers, real-ear probe-microphone systems, and hearing aid fitting software, and thus they provide a primary means to select and verify the frequency-gain characteristics of hearing aids. Aside from its use in hearing aid selection and verification (Popelka and Mason, 1987; Pavlovic, 1988; Rankovic, 1991; Mueller, 1992; Studebaker, 1992; Studebaker and Sherbecoe, 1993), the Articulation Index has also been used to differentiate between sensory and neural auditory lesions (Gates and Popelka, 1992) and to predict speech intelligibility of listeners who wear personal hearing-protection devices (Wilde and Humes, 1990).

Because the term *articulation* suggests the act of speech production rather than audition, it is misleading in today's vernacular. Thus it has been suggested that *articulation* be substituted by *audibility,* and reference to this index be termed *Audibility Index* (Studebaker, 1992; Killion *et al.*, 1993; Killion, 2002). Despite this recommendation, the American National Standards Institute (ANSI) adopted *Speech Intelligibility Index,* in recognition of the critical concept of predicting speech intelligibility (ANSI S3.5-1997).

To minimize confusion, Articulation Index will be used in this paper to refer to the principles and procedures described in the original ANSI S3.5-1969 standard, measures derived directly from this procedure, and a set of general methods in which articulation theory is applied. The Speech Intelligibility Index (SII), on the other hand, will refer to those principles and procedures described in the most recent standard (ANSI S3.5-1997) or to derivative measures of this newer method. Because the Articulation Index and Speech Intelligibility Index are based, in essence, on determining the amount of audibility available to a listener, we refer to these indices collectively as the Audibility Index throughout the remainder of the paper. Incidentally, to avoid introducing a new abbreviation, we will use the abbreviation AI to refer interchangeably to the Audibility Index and the Articulation Index.

Based on its clinical potential and the amount of ongoing research aimed at improving its predictive validity, the AI promises to play a significant role in the practice of audiology for many years to come. It behooves audiologists, therefore, to understand the relevant concepts. Determining the AI has been viewed as a complicated and confusing procedure. In this paper, we take on the decidedly difficult task of providing a somewhat simplified, yet detailed, explanation of articulation theory as it relates to clinical audiologic practice. While mathematical formulas are necessarily a part of any such description, our goal is to provide sufficient narrative explanation of the formulas to give even mathematically challenged readers a general understanding of the various methods and applications.

This paper is divided into three major sections. In the first section, the fundamental principles and calculation methods of the AI as specified in the ANSI S3.5-1969 and ANSI S3.5-1997 standards are described, along with clinical simplifications of these procedures and other related methods used to establish the audibility of speech under diverse conditions. The second section provides a review of the literature on the use of the AI to predict speech intelligibility for individuals with various degrees of hearing impairment. The third section presents a discussion about the use of the AI to predict speech intelligibility in clinical hearing aid applications.

Principles and Calculation Methods

A. Articulation Index (ANSI S3.5-1969 Standard)

Descriptions of the basic model and calculation method of the AI (French and Steinberg, 1947) were based on a compendium of studies that aimed to determine the factors affecting speech intelligibility. These included the ideal or optimal speech spectrum (Dunn and White, 1940); the effects of masking (Fletcher and Munson, 1937); the frequency of steady-state noise and the intensities of noises from several sources, including ambient noise (Waring, 1946; French and Steinberg, 1947); zero loudness-contour curves for one and two ears (ASA Z24.2, 1942); and the number of frequency bands needed to estimate speech intelligibility (French and Steinberg, 1947).

Over the next 15 years, investigators amended and modified the AI. These addenda, however, were not generally accepted because of insufficient evidence of their validity (Kryter, 1962[b]). As a means to validate the AI, Kryter (1962[a], 1962[b]) published a series of papers that resulted in the ANSI S3.5-1969 standard.

According to the original standard, the AI is a proportional index based simply on the summed audibility of weighted speech bands in quiet and in the presence of competing noise measured at a listener's ear. It ranges in value from 0.0 to 1.0. A value of 0.0 suggests that none of the speech cues are audible, thereby making zero contribution to speech intelligibility. A value of 1.0, on the other hand, indicates that the proportion of available, or audible, speech cues contributes maximally to speech intelligibility. For values between 0.0 and 1.0, a proportionate growth of audibility is assumed to occur, and with it, a proportionate growth in intelligibility performance. That is, within any frequency band, every decibel of audibility is weighted equally.

The basic equation for the Articulation Index is

$$AI = \Sigma I_i A_i \quad (1)$$

In the formula, AI is equal to the sum of products resulting from multiplying I by A at each frequency band (i). The frequency-importance function (I_i) represents the relative contribution of different frequency bands (ie, fractions of 1.0) to speech intelligibility. The variable A_i, or audibility function, which ranges from 0 to 30 dB, refers to the amount of speech energy that is above the listener's threshold and any competing noise in a given frequency band. Before illustrating the calculation of the AI, we will first review how the frequency-importance function, or I_i is derived.

Historically, the number of bands used to derive the frequency-importance function has varied from procedure to procedure, ranging from 4 to 21 (ANSI S3.5-1997). Initially, French and Steinberg (1947) defined the importance function by dividing the frequency spectrum into 20 octave bands ranging in frequency from 250 to 7000 Hz, such that each band made an equal contribution (0.05) to speech intelligibility. This concept was adopted in the original (ANSI S3.5-1969) standard and modified to incorporate the addition of one-third octave bands. Of these two procedures, the one-third octave-band method is known to be more sensitive to variations in speech and noise (Kryter, 1962[a]; ANSI S3.5-1969).

The value of I_i is some proportion between 0.0 and 1.0, and it is chosen so that AI = 1.0 when A_i = 30. The values are consistent with the number of bands and bandwidth used in the derivation, and the relative importance of each frequency band varies with the speech material. As shown in Figure 1, the one-third octave bands

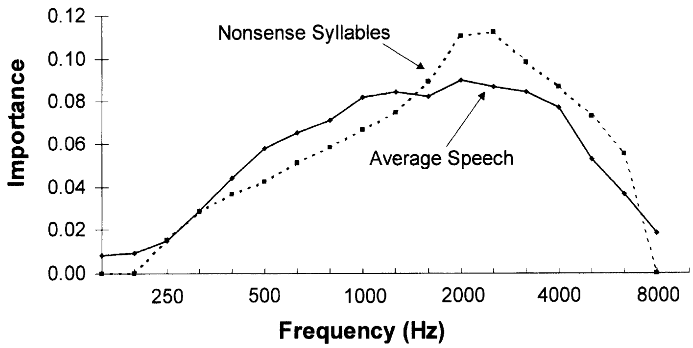

Figure 1. Importance functions for average speech (solid line) (Pavlovic, 1987) and nonsense syllables (dashed line) (French and Steinberg, 1947).

centered around 2000 Hz are most important for the recognition of nonsense syllables, whereas the one-third octave bands ranging from 800 to 4000 Hz are of roughly equal importance for understanding average speech.

The mathematical weightings that comprise the frequency-importance function are determined from the results of speech-intelligibility tests presented under various filtering conditions. For each filter condition, the maximum amount of speech information carried by a given band is determined.

French and Steinberg (1947) examined the effects of various high-pass and low-pass filter cutoff frequencies on the intelligibility of nonsense syllables (Figure 2). They did this using consonant-vowel (CV), vowel-consonant (VC), and consonant-vowel-consonant (CVC) nonsense syllables created at Bell Telephone Laboratories and spoken by roughly an equal number of male and female talkers.

Figure 2 shows the percent-correct scores achieved by a group of normal-hearing listeners when listening to filtered nonsense syllables as a function of filter cutoff frequencies. Results demonstrated that as bandwidth decreased, speech-intelligibility scores also decreased for both filter types. For the sets of paired filter conditions studied, the frequency (A) for the high-pass and low-pass conditions was regarded as the midpoint, or point at which intelligibility was equal on either side. At this crossover frequency, the percent intelligibility was less than 100%, but greater than 50% in each band. The crossover frequency for syllables indicated that an equal amount of intelligibility (50%) was carried at frequencies below and above 1900 Hz.

In 1947, Beranek proposed a modified frequency-importance function for nonsense syllables based on male talkers only. This change, advocated by Kryter (1962[a], 1962[b]), was adopted in the ANSI S3.5-1969 standard. It should be noted that other investigators also provided new frequency-importance functions (Fletcher and Galt, 1950; Black, 1959), but these functions were fairly similar to those derived by French and Steinberg (1947) and were not incorporated into the original standard.

The second variable of the AI formula, A_i, designates how much of the available information given by I_i is actually delivered to the listener. The dynamic range of speech that is important for intelligibility is assumed to be 30 dB. In other words, each decibel above threshold represents 1/30 of the range of audible signal that contributes to speech intelligibility within a given frequency band.

The value of A_i can be calculated for quiet conditions by subtracting a given listener's threshold from the speech maxima of the idealized long-term average speech spectrum. This calculation is performed for each individual frequency band. The effective noise level includes all ambient

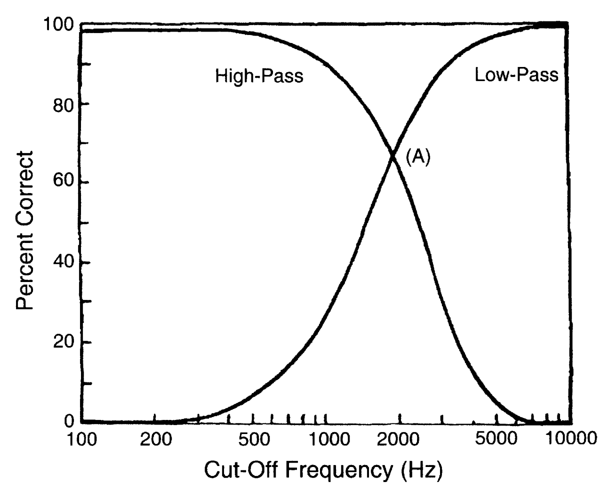

Figure 2. Correct identification of syllables as a function of high-pass and low-pass filtering. Adapted from French and Steinberg. Factors governing the intelligibility of speech sounds, J Acoust Soc Am 19:90-119, 1947. Reprinted with permission of the Acoustical Society of America.

noise, as well as the hearing threshold level, which is converted to a level of fictitious internal noise (French and Steinberg, 1947).

As Figure 3 illustrates, the contribution of a given frequency band is 0.0 when speech is below an individual listener's hearing threshold (or when the noise level exceeds the speech spectrum level). When the entire 30-dB range of speech is above the listener's hearing threshold, the band makes maximal contribution to speech intelligibility (ie, 1.0). The AI, therefore, is the sum of the weighted audibility across all frequency bands.

The use of the idealized speech peaks in the AI calculation suggests that accurate results can best be obtained if it is assumed that the short-term speech distribution is uniform over time. Specifically, the speech spectrum of the CVC syllables of the Bell Telephone Laboratories was used as the basis for the original AI (ANSI S3.5-1969) standard.

Dunn and White (1940) used the average sound pressure level within a band 1 Hz wide, over the integration time of the ear (ie, 125 milliseconds). The accuracy of predictions of speech intelligibility over an integration time of 125 milliseconds has since been verified in other studies (French and Steinberg, 1947; Kryter, 1962[b]; Pavlovic and Studebaker, 1984). Dunn and White found that within this integration time, the distribution of the speech root-mean-square (RMS) values is approximately linear over a 30-dB range in any given frequency band. This 30-dB range extends from +12 dB to -18 dB between 1000 and 1400 Hz relative to the long-term average speech spectrum of the CVC syllables in the original standard.

This range was modified to ±15 dB in the revision of the ANSI S3.5-1969 standard, which was drafted as the ANSI S3.5-1993 interim standard (ANSI, 1993). The dynamic range of ±15 dB in the ANSI S3.5-1997 standard resulted from speech-recognition studies reflecting differing speech stimuli in speech-weighted noise at various signal-to-noise ratios (SNRs) (eg, Studebaker *et al.*, 1993) and the methods used in the calculation of the Speech Transmission Index, which is described later in this paper.

Thus, for CVC syllables, the overall RMS speech level minimally required for recognition of weak consonants is 30 dB higher than that required minimally for recognition of strong vowels. This range has also been termed "perceptual dynamic range" (Boothroyd, 1986, as referenced in Pavlovic, 1987 and 1989).

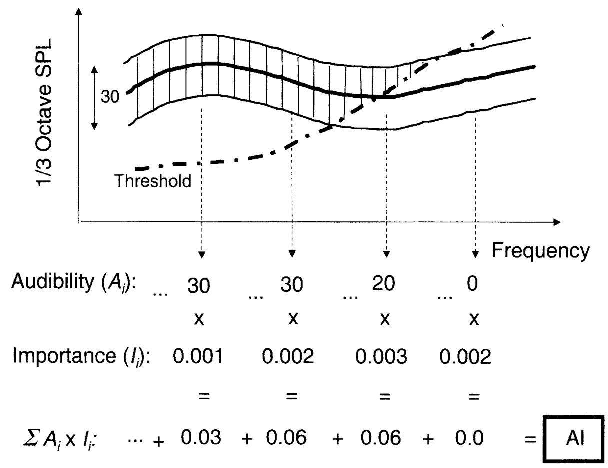

Figure 3. An illustration of audibility and band importance as a function of threshold in the calculation of the AI.

The idealized speech spectrum is assumed to be based on speech stimuli presented in quiet, with normal vocal effort, and at a distance of 1 meter from the talker. The overall presentation level of speech produced by normal vocal effort is further assumed to be 65 dB SPL in the original AI standard and 63 dB SPL for the SII,[1] as measured in a sound field. Under everyday listening conditions, however, various factors may change the amplitude characteristics of speech. For instance, if a high level of background noise is present, a talker is more likely to raise the voice level due to the Lombard effect. That is, for every decibel increase in noise above 50 dBA, there is an increase in speech level of 0.46 dB (ANSI S3.14-1977). Pearson *et al.* (1976) found that increased vocal effort is associated not only with higher overall level, but also with variations in the frequency-amplitude spectrum. A second factor to consider clinically is that a given hearing-impaired listener may participate in conversational speech at listener-to-talker distances that vary considerably in different conversational situations.

[1]The AI accounts for variations in vocal effort using a correction factor (Figure 13, ANSI S3.5-1969). Conversely, the SII assumes values of 69, 75, and 82 dB SPL for raised, loud, and shouted vocal efforts, respectively.

The following conclusions can be drawn regarding the ANSI S3.5-1969 standard:

1. The AI is based on two components: the importance of each frequency band (I_i) and the amount of signal audibility within a given band (A_i).
2. The frequency-importance (I_i) function is a mathematical weighting that represents the importance of different frequencies for speech intelligibility. The importance of these weightings differs with respect to the speech material, and the sum of these weightings must equal 1.0.
3. The audibility (A_i) function quantifies how much of the signal within a given frequency band is being delivered to the listener. When the idealized speech spectrum is used in the calculation, this function assumes normal vocal effort and a conversational level of 65 dB SPL in the original standard and 63 dB SPL in the newer standard, at a distance of 1 meter from the talker. Changes in these levels, whether related to hearing threshold, competing noise, or listener-to-talker differences, will result in concomitant changes in signal audibility for the listener.
4. The dynamic range of speech that is regarded to be effective in maximizing speech intelligibility is 30 dB. This range was originally found to be between +12 and -18 dB, and was recently modified to ±15 dB.

B. Calculation of the AI

To calculate the AI, the audiometric thresholds of a listener are converted from dB HL to an equivalent level of a fictitious internal noise in the listener's ear. This is done by first converting dB HLs to the reference equivalent threshold sound pressure levels. An illustrative example of the calculation of the AI is shown in Table 1, which uses transformations from Pavlovic (1987) and Bentler and Pavlovic (1989) to convert HLs of a hypothetical hearing-impaired listener tested with TDH-49 earphones to equivalent SPLs in free

Table 1. *Example of AI Calculation Using Hypothetical Audiometric Data for a Hearing-Impaired Listener Tested with TDH-49 Earphones*

	Frequency (Hz)				
	250	500	1000	2000	4000
Hearing thresholds (dB HL)	10	20	30	40	50
+ Conversion to dB SPL in free field[1]	12.7	7.5	5.7	2.5	-1.9
- Critical ratios[2]	16.6	17.2	18.2	20.2	24
+ 10 $\log_{10}$(bandwidth)[3]	22.5	25.5	28.5	31.5	34.5
= Equivalent threshold	28.6	35.8	46	53.8	58.6
Idealized speech peaks for octave bands[4]	72.5	74	68	62	57
A_i (audibility function)	30*	30*	22	8.2	0**
x I_i (importance function)[5]	0.0024	0.0048	0.0074	0.0109	0.0078
A_iI_I	0.072	0.144	0.163	0.089	0
			AI = 0.468		

[1]Bentler and Pavlovic (1989); [2]Pavlovic (1987); [3]ANSI S3.5-1997, Table 4; [4]ANSI S3.5-1969, Table 8; [5]ANSI S3.5-1969, Table 7. These values are based on nonsense syllables developed at Bell Telephone Laboratories.

*When this value is greater than 30, the assigned value is terminated at 30.
**When this value is less than 0, the assigned value is terminated at 0.

field. The equivalent level of the fictitious noise is taken to be equal to the band pressure level minus the critical ratio plus 10 $\log_{10}$ (bandwidth). In each band, the idealized speech spectrum maxima are compared to the equivalent thresholds. When the speech maximum is at threshold, the signal in that band makes zero contribution to intelligibility. When the speech maximum is 30 dB or more above the threshold, it makes maximal contribution to intelligibility. In each band, the amount of audible signal is weighted by the relative importance of that frequency band for speech intelligibility. The sum of weighted audibility across all bands is the AI, which in this example is 0.47. (The adequacy of applying the AI model that was originally developed for normal-hearing listeners to hearing-impaired listeners will be addressed in the next major section).

C. Using the AI to Estimate Speech Intelligibility

Once a predicted value has been calculated for a given listener, it can be used to estimate the intelligibility of various kinds of speech stimuli by means of a transfer function. Initially, Fletcher and Galt (1950) described this facet by monotonically relating speech intelligibility, in percent correct, to AI for normal-hearing listeners. This relationship is defined by the power function

$$S = (1-10^{-AP/Q})^N \quad (2)$$

where S is the score in proportion correct, A is the AI value, and P is a proficiency factor, which accounts for variables relating to practice and experience of the talker and the listener (Studebaker and Sherbecoe, 1993). Q and N are fitting constants, the values of which depend on the speech stimulus and the subjects tested.

To simplify the process of predicting speech intelligibility from equations, a graphical articulation-to-intelligibility transfer function is also available (Webster, 1979). As seen in Figure 4, a normal-hearing listener with an AI of .60 would be expected to score approximately 75% on lists of nonsense syllables, 80% on lists of phonetically balanced words, and 98% on sentences. It should be noted that measured speech scores would differ somewhat from the predicted scores based on the AI, depending on the actual speech stimulus used and the proficiency of the talker and listener.

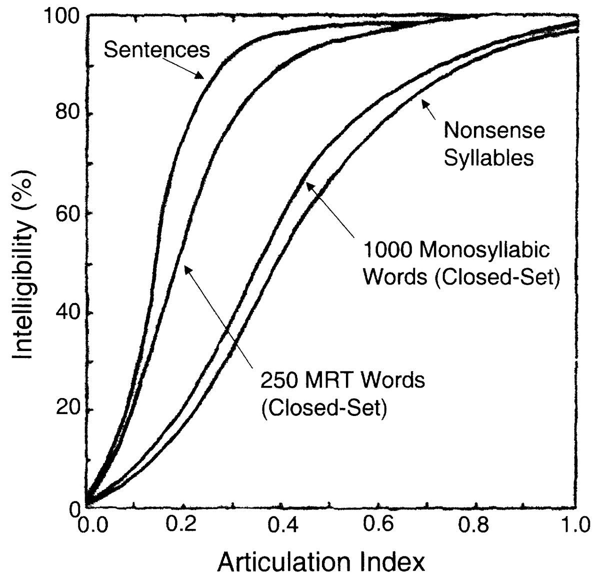

Figure 4. Speech intelligibility correct, in percent, for various speech stimuli as a function of articulation index. Adapted from Webster, Interpretations of speech and noise characteristics of NTID learning centers, J Acoust Soc Am 66S:S37, 1979, as reported in Killion (1985). Reprinted with permission of the Acoustical Society of America.

D. Simplified Methods of Calculating AI

There have been several attempts at simplifying the calculation of the AI for clinical use (Pavlovic, 1988, 1991; Mueller and Killion, 1990; Humes, 1991; Kringlebotn, 1999). All of these schemes make use of an audiogram display, so that the hearing thresholds can be plotted directly onto the audiogram for calculating the AI. Here, we provide the reader with detailed descriptions of these simplified methods. The differences between these methods are summarized in Table 2, and examples using the different simplified methods for calculating the AI for the same hypothetical hearing-impaired listener as seen in Table 1 are provided in Figures 5–10. Note that the same unaided and aided thresholds are used in these examples.

Despite similarities in their conception, each procedure differs with regard to the amount of frequency weighting used. Pavlovic (1988), who developed the precursor to the present-day meth-

Table 2. *Simplified Methods for Calculating the AI*

Authors	Frequency Bands (Hz)	Importance Function	Dynamic Range	Calculation of the AI
Pavlovic (1988) $A_0(4)$	500, 1000, 2000, 4000	None (equal weighting)	30 dB represented by shaded area	AI = Sum of audible decibels at each frequency/120
Mueller and Killion (1990)	250, 500, 1000, 2000, 3000, 4000, 6000	Nonsense syllables, represented by 100 dots	30 dB, standard speech spectrum	AI = Number of dots below threshold curve × 0.01
Humes (1991)	250, 500, 1000, 2000, 4000	Nonsense syllables, represented by 33 dots	Varies between 30–40 dB across frequencies	AI = Number of dots below threshold curve × 0.03
Pavlovic (1991), Lundeen (1996)	250, 500, 1000, 2000, 3000, 4000, 6000	Average speech, represented by 100 dots	30 dB	AI = Number of dots below threshold curve × 0.01
Pavlovic (1991) $A_0(6)$	250, 500, 1000, 2000, 3000,4000, 6000	None (equal weighting)	30 dB	AI = (Sum of audible decibels between 500–2000 Hz+ *average* audible decibels between 3000–6000 Hz) /120
Kringlebotn (1999)	250, 500, 1000, 2000, 3000,4000, 6000	Monosyllabic words, represented by 100 dots	30 dB	AI = Number of dots above threshold curve × 0.01

ods, suggested that hearing levels ranging between 20 and 50 dB HL and for the frequencies of 500, 1000, 2000, and 4000 Hz form the boundaries for all speech cues important to speech recognition. In his original method, termed A_0 and later renamed $A_0(4)$, clinicians can easily calculate an AI for the audible portion of the dynamic range of speech (ie, 30 dB) across the four frequencies using a standard audiogram. As seen in Figure 5, our hypothetical patient exhibits essentially a mild-to-moderate sloping hearing loss in the right ear. At 500 Hz, a threshold of 20 dB HL is observed. Therefore, all 30 dB of speech energy around 500 Hz is considered audible. For the frequencies of 1000 and 2000 Hz, unaided thresholds are measured at 30 and 40 dB HL, respectively. The number of audible decibels is 20 at 1000 Hz and 10 at 2000 Hz. Note that at 4000 Hz, the threshold is 50 dB, resulting in 0 dB of audibility. To determine the amount of available speech information, one simply sums the number of decibels audible to the listener, which in this case is 60 dB (30+20+10+0), and divides the total by 120, or the maximum number of audible decibels (30 dB × 4 frequencies). This results in an unaided AI of 0.50 for this patient's right ear, suggesting that exactly one half of the speech spectrum is audible.

Suppose this individual is fit with a hearing aid and sound-field thresholds are measured either by traditional audiometric methods or derived by using real-ear aided gain (REAG) values obtained using probe-microphone techniques, resulting in aided thresholds of 10, 15, 20, and 30 dB for the frequencies between 500 and 4000 Hz. For those respective frequencies, this results in aided gain values of 10, 15, 20, and 20 dB. A recalculation of the aided values using Pavlovic's (1988) $A_0(4)$ method results in 110 audible decibels out of a possible 120 decibels. This equates to an aided AI of 0.92 (110/120), a 42% increase in the speech cues now afforded the listener with amplification. Note that this method is equally weighted at each of the four audiometric frequencies and that the audiogram in Figure 5 does not conform to the ANSI S3.21-1978 (R1997) specifications, where one octave on the X-axis should equal 20 dB on the Y-axis.

In response to Pavlovic's (1988) $A_0(4)$ procedure, Mueller and Killion (1990) devised a procedure to account for the inter-octave frequencies of 3000 Hz and 6000 Hz, as well as to simplify the denominator during the division process. To account for these changes, Mueller and Killion used the count-the-dot method originally described by Cavanaugh *et al.* (1962) and altered the importance function based on nonsense syllables and one-third octave frequency bands, including the inter-octave frequencies of 3000 Hz and 6000 Hz.[2] For clinical implementation, this count-the-dot method was simplified to 100 dots and superimposed on an audiogram (Figure 6). Again, the number of dots occurring at a specific frequency corresponds to frequency-importance weightings, whereas the proportion of energy occurring above threshold at a given frequency (physically under the threshold curve on the audiogram) represents the proportion of energy that is audible in a particular frequency-importance band (ie, effectively, A_i).

To calculate an AI using this procedure, therefore, one need only count the number of dots audible to the listener (the number of dots falling physically below, or touching, the threshold curve) and multiply by 0.01. Using thresholds of 10 to 60 dB HL between the audiometric frequencies of 250 to 8000 Hz, 47 dots of audibility are available, resulting in an unaided AI of 0.47. Assuming aided thresholds ranging in level from 5 to 45 dB HL for the same audiometric frequencies, 0.90 of the speech cues are now available to

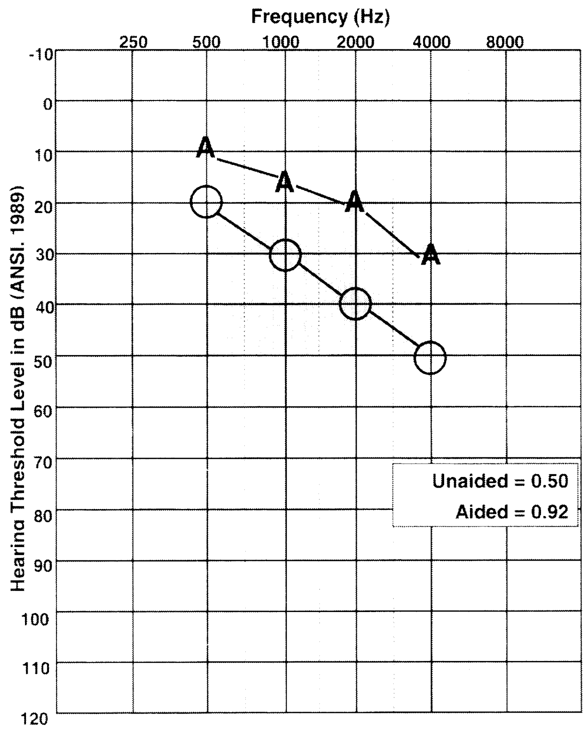

Figure 5. Example of Pavlovic's (1988) $A_0(4)$ procedure for the unaided (circle) and aided (A) conditions.

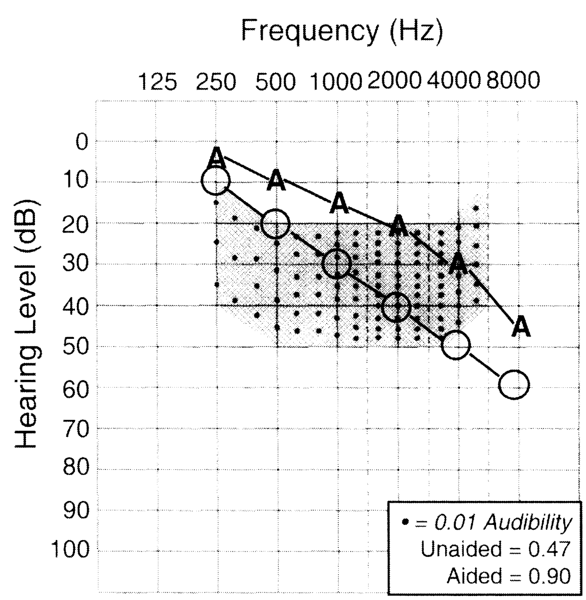

Figure 6. Count-the-dot-audiogram by Mueller and Killion (1990) for the unaided (circle) and aided (A) conditions. Adapted from Killion *et al.*, A is for audibility, Hear J 46(9):29-32, 1993. Reprinted with permission of Lippincott Williams & Wilkins.

[2] A 200-dot SPL-O-Gram, developed by Webster in 1979, can also be considered a precursor to contemporary count-the-dot audiograms.

the listener. As with any other method designed to take inter-octave frequencies into account, thresholds at these frequencies must be measured or interpolated from audiometric data.

For aided conditions, the impact of output limiting must also be considered. Measurement of uncomfortable loudness levels (UCLs) plays an important role in calculating the AI. In other words, the AI is calculated based on the number of dots (in the case of the count-the-dot methods), or audible decibels at each frequency between aided threshold and aided UCL. When speech cues are deemed to be above the aided UCL, they are assumed to be limited by the hearing aid and therefore inaudible. Using Figure 5 as an example, assume that an unaided UCL of 90 dB HL is measured for the audiometric frequency of 1000 Hz. Based on the 15 dB of gain provided by the hearing aid, the aided UCL is assumed to be 75 dB HL, or the unaided UCL minus the amount of gain (90 – 15). Here, the aided UCL would not impact the AI calculation. If the unaided UCL were 60 dB HL, as seen potentially with a hyperacusic patient, the derived AI would indeed be affected. The examples presented here are relevant to linear signal processing. In cases where nonlinear signal processing is used, the gain values determined at the *high input levels* should be applied when calculating aided UCL values.

As illustrated in Figure 7, Humes (1991) proposed a revision to the Mueller and Killion (1990) method. Specifically, Humes' count-the-dot procedure differs from Mueller and Killion's (1990) in that (1) it uses an octave-band resolution, as opposed to a one-third octave-band resolution; (2) the dynamic range of speech varies between 30 and 40 dB across frequencies, based on the work of Pascoe (1980); (3) the importance functions are derived from the work of French and Steinberg (1947), resulting in higher importance weightings in the mid and high frequencies; and (4) the number of dots is reduced to 33, making counting easier. As a result of these modifications, a greater number of dots occur at the audiometric frequencies of 1000, 2000, and 4000 Hz than at other frequencies. Counting the number of dots available to the listener and multiplying that value by 0.03 determines audibility in this procedure. For a listener demonstrating the audiometric thresholds described earlier, 17 dots would fall above the listener's threshold (below the threshold curve), yielding an unaided AI of 0.51 (17 × 0.03). Notice that if the listener is fit with an amplification device and real-ear insertion gain (REIG) is measured, thresholds are 5, 10, 15, 20, and 30 dB at the frequencies between 250 and 4000 Hz, and the number of dots below the aided curve increases to 30, resulting in an aided AI of 0.90.

In 1991, Pavlovic concluded that the Mueller and Killion (1990) and Humes (1991) methods might not sufficiently predict performance in everyday listening situations. In essence, his viewpoint was that an importance function for average speech is more appropriate for AI calculations derived with hearing aids. Termed the A_d method (the d being a mnemonic for "dot"), this count-the-dot procedure is similar to Mueller and Killion's (1990) method in every respect (including calculation), except that the speech spectrum used is that of conversational speech in quiet, measured at 63 dB SPL and at a distance of 1 meter from the speaker. Thus the A_d method is applicable directly to continuous discourse.

Lundeen (1996) noted the need to make two corrections in the published version of Pavlovic's A_d method (1991). Specifically, the physical width of an octave on the frequency scale of the audiogram (X-axis) did not equal the length of a 20-dB interval on the ordinate, or Y-axis, as specified by ANSI S3.21-1978 (R1997). Also, the dots

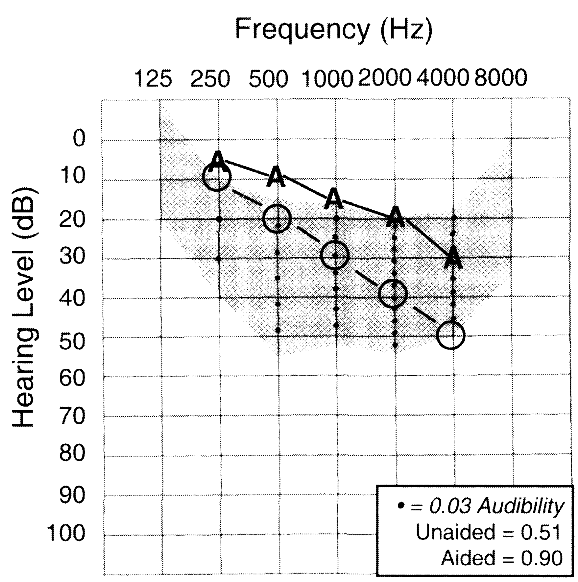

Figure 7. Count-the-dot-audiogram by Humes (1991) for the unaided (circles) and aided (A) conditions. Adapted from Killion *et al.*, A is for audibility, Hear J 46(9):29-32, 1993. Reprinted with permission of Lippincott Williams & Wilkins.

above 4000 Hz were spaced at different frequency intervals relative to all other frequencies on the audiogram. Lundeen's modified A_d procedure is shown in Figure 8. Using the same audiometric hearing loss described throughout this subsection, in addition to the threshold of 60 dB at 8000 Hz, an AI value of 0.47 is derived based on multiplying 0.01 by the 47 dots determined to be below the threshold curve. After amplification, the number of dots available to the listener below the aided curve has increased by 36, resulting in an aided AI of 0.83.

Pavlovic (1991) also modified the original $A_0(4)$ method, but only as a means to account for the frequencies of 3000 and 6000 Hz. This new method, termed $A_0(6)$, is calculated differently than its predecessor and includes four steps. To understand its use, one can use the same thresholds (20, 30, 40, and 50 dB HL for the respective frequencies of 500, 1000, 2000, and 4000 Hz) as used in the earlier-described $A_0(4)$ procedure (Figure 5), and assume thresholds of 45 and 55 dB HL at 3000 and 6000 Hz, respectively (Figure 9). First, a sum of the audible decibels from 500, 1000, and 2000 Hz is obtained. This value is 60 (30 + 20 + 10). Next, the sum is determined for the higher three frequencies, resulting in a value of 5 (5 + 0 + 0). The third step calls for totaling the number of audible decibels. This is done by dividing the results from the second step (5) by the value 3 and adding the newly derived value (1.7) to the audible decibels determined in the first step (60). This results in a value of 61.7. Using this value of 61.7 as a numerator, the final step requires dividing by the denominator of 120 (61.7/120), which results in an unaided AI of .51. Using the aided thresholds shown in Figure 9, the calculation is repeated (110/120) to derive the aided AI, which is 0.92. Note that both the unaided and aided values are weighted equally. (Once again, the audiogram in Figure 9 does not conform to the ANSI S3.21-1978 [R1997] specifications, where 1 octave on the X-axis should equal 20 dB on the Y-axis.)

Kringlebotn (1999) recently derived a graphical method based on the newer standard for calculating the AI from audiometric data. In its simplest form, this model assumes monaural listening at a distance of 1 meter from a talker who is expending normal vocal effort. The speech area, defined for CID-W22 monosyllabic words (Hirsh *et al.*, 1952), consists of 100 points (10 rows of 10 dots) transposed onto an audiogram (Figure 10).

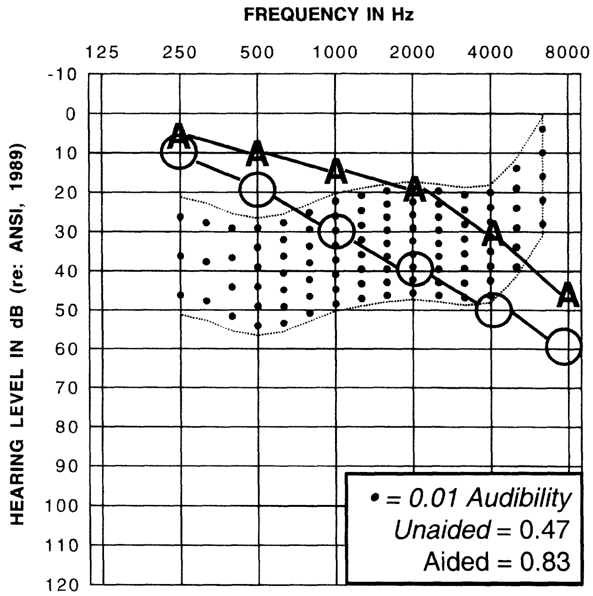

Figure 8. A modified version of Pavlovic's (1991) A_d count-the-dot-audiogram for the unaided (circle) and aided (A) conditions. Adopted from Lundeen, Count-the-dot audiogram in perspective, Am J Audiol 5:57-58, 1996. ©American Speech-Language-Hearing Association. Reprinted with permission.

Thus, the contribution of each dot to audibility is 0.01. Notice that the hearing levels are reversed with respect to the traditional audiogram, requiring dots above the threshold curve (as well as above threshold) to be counted. Assuming the same audiometric thresholds for the octave frequencies between 250 and 4000 Hz described earlier, and adding a 60 dB threshold at 6000 Hz, our hypothetical patient's audiometric values reveal 52 dots to be audible. AI would thus equate to 0.52 (52 × 0.01), signifying the availability of slightly more than one half of monosyllabic-word cues. This model can also be used to predict the AI after amplification produced by a linear hearing aid by including a correction factor for desensitization, or the reduction in the ability of a sensorineural impaired ear to extract audible information that contributes to speech intelligibility (Pavlovic *et al.*, 1986). Clinically, however, the derivation of this corrected AI is not simple, and requires the use of a calculator or personal computer.

Results from the above examples, which used the same audiometric thresholds, yielded AI val-

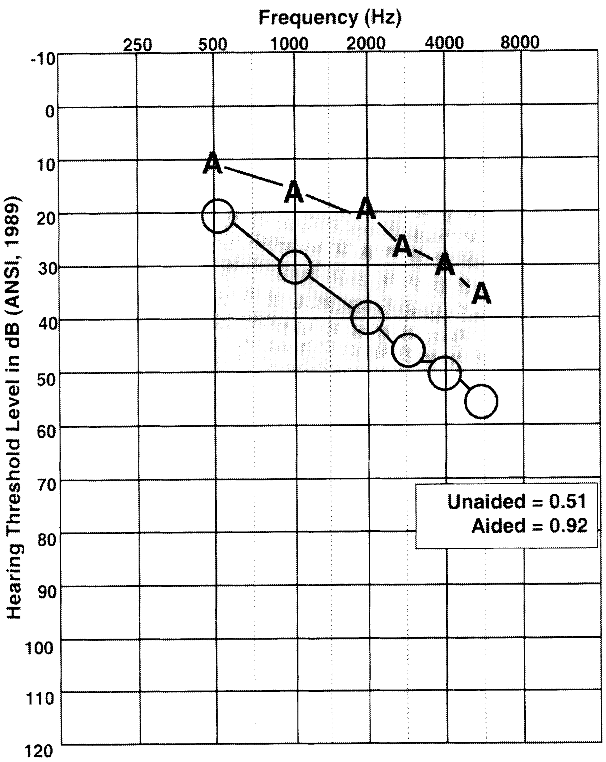

Figure 9. Example of Pavlovic's (1988) A_0 (6) procedure for the unaided (circle) and aided (A) conditions.

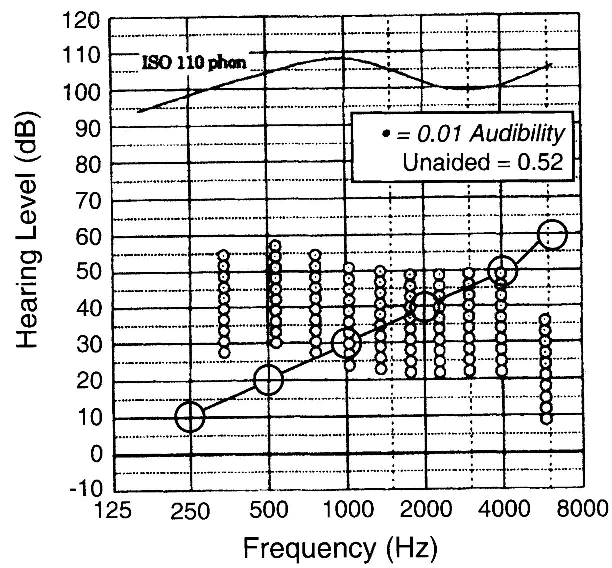

Figure 10. Count-the-dot-audiogram by Kringlebotn (1999) for the unaided (circle) condition only. (Reprinted with permission from the International Journal of Audiology)

ues ranging from 0.47 to 0.52 when unaided and .90 to .92 for all but one procedure (Lundeen, 1996) when aided. The different count-the-dot methods, therefore, result in remarkably similar outcomes. Few studies in the literature, however, have assessed the accuracy of the various simplified methods. Humes and Riker (1992) evaluated the predictive accuracy of Pavlovic's (1988) $A_0(4)$ method and Humes' (1991) count-the-dot methods. Unaided AI calculations were derived for 10 normal-hearing young adults (19 to 29 years) and unaided and aided measures derived for 14 hearing-impaired elderly adults aged 65 to 75 years. Northwestern University Number 6 (NU-6) (Tillman and Carhart, 1966) monosyllabic words were presented in quiet at an SNR of +7 dB in sound field, with both speech and noise stimuli presented from directly in front (0° azimuth). Results indicated that both methods yielded similar values, particularly for the unaided condition in quiet, for the young and elderly groups. For the elderly adult group, however, aided speech-recognition performance was less predictable by both methods, particularly in the presence of noise. Humes (1991) found similar aided results using younger listeners, nonsense syllables, and a quiet listening condition.

Similarly, Souza *et al.* (2000) compared the relationship between the AI values obtained via real-ear analyzer software and the traditional AI for 115 hearing-impaired listeners. The AI value was based on the $A_0(4)$ procedure developed by Pavlovic (1988) and modified to incorporate conversational speech levels (Pavlovic, 1991). Although it is not clear from the published report which variables (eg, aided, target) were being compared, these investigators found a fairly strong relationship ($r \geq 0.86$) for both ears between the clinical and traditional versions of AI. Based on their findings, Humes and Riker (1992) and Souza *et al.* (2000) recommended that Pavlovic's (1988) $A_0(4)$ method (Figure 5) be implemented clinically due to its simplicity. Note that there is no need to count dots with this method, but merely to note the number of decibels (in 5-dB steps) below the threshold curve.

The use of these simplified AI methods has both benefits and limitations. First, incorporating the audiogram has considerable value in demonstrating to patients the effects of hearing impairment on speech intelligibility without and with the use of amplification. A second distinct advantage is that these methods enable clinicians to describe the extent to which speech cues are available, independently of the need to perform speech-intelligibility testing. In addition, count-the-dot audiograms provide a means whereby different hearing aids and types of circuits can be compared as a function of the speech input level.

While these methods are a simplified, effective, and efficient means of estimating the proportion of speech cues available to a given listener, clinicians must be aware of several factors that could yield inaccurate results. Some but not all of these inaccuracies stem from shortcomings inherent in the AI itself. Clinicians, for example, may mistake audibility for intelligibility. As described earlier in this paper, audibility and intelligibility are related, but they are not synonymous. For instance, 0.6 audibility does not equate to 60% intelligibility, but rather indicates only that 60% of the speech cues are available to the listener. The prediction of speech intelligibility can be determined based on the transfer functions for various speech stimuli, as seen in Figure 4, for example. While speech-intelligibility performance is assumed to increase monotonically with increasing sensation level or audibility, this is generally true only for those individuals with normal hearing or with mild-to-moderate hearing losses. As discussed elsewhere in this paper, for individuals with more severe hearing losses, an increase in sensation level (ie, audibility) may not result in improved intelligibility, but may actually degrade intelligibility.

Pavlovic (1991) further warns that these simplifications might mislead clinicians and patients in several other ways. First, a value of 1.0 does not necessarily mean that the aided auditory system is functioning normally. Rather, it indicates that the hearing aid is matched optimally to the impaired auditory system for maximizing speech intelligibility. Second, improved thresholds, as measured by REIG, may suggest that the listener's threshold has changed when, in actuality, it is the speech spectrum that has been shifted to a more audible range. Lastly, and perhaps most importantly, these simplified AI methods cannot predict a patient's ability to understand speech in noise (Killion and Christensen, 1998; Killion, 2002).

E. Speech Intelligibility Index (ANSI S3.5-1997 Standard)

In the ANSI S3.5-1997 standard, the term Articulation Index was changed to Speech Intelligibility Index, presumably to focus on the SII's objective of speech intelligibility prediction. As in the previous standard, audibility is expressed by the equation

$$SII = \Sigma I_i A_i \tag{3}$$

where I_i and A_i are the frequency-important and audibility functions, respectively.

Several modifications and additions to the original ANSI S3.5-1969 standard characterize the 1997 standard, and these are detailed in this subsection. In the ANSI S3.5-1997 standard, for instance, the calculation of the audibility function has been modified to consider the spread of masking and the standard speech spectrum level. That is, because the world is a noisy place, target sounds listeners want to hear are mixed with environmental, background noises. When speech is mixed with noise, some parts of the speech spectrum become inaudible due to masking—a phenomenon with which audiologists are very familiar. Masking is also problematic with respect to speech intelligibility when speech masks itself, or when higher energy vowels make lower energy consonants inaudible. Empirical evidence suggests that a reduction in consonant information can yield decreased speech-intelligibility performance because about 90% of the acoustic information important for understanding is provided by consonants (French and Steinberg, 1947; Licklider and Miller, 1951). The effect of masking depends on the various parameters of noise: (a) its long-term spectrum, (b) its intensity fluctuation over time, and (c) its average intensity relative to the intensity of speech.

To account for masking, the 1997 standard includes formulas for self-speech masking and upward spread of masking. These variables are referenced and detailed in Appendix A. The revised calculation method also provides a framework for determining the speech, noise, and threshold spectrum levels based on measurements of modulation transfer functions in reverberation.

In addition, the revised standard takes into consideration the empirical findings of French and Steinberg (1947) and Fletcher (1953) that suggested a decrease in speech performance for normal-hearing listeners at high sound pressure levels (ie, rollover). As a means to counter the phenomenon of rollover, the audibility function has been modified to include a level distortion factor (LDF). This factor has a maximum of 1 when there is believed to be no distortion, typically at low presentation levels. This factor, however, is reduced when the overall sound pressure level exceeds 73 dB SPL, and reaches a minimum value of 0 only at extremely high presentation levels (ie, 233 dB SPL). The modified audibility function A_i is expressed as

$$A_i = K_i L_i \tag{4}$$

where K_i is the proportion of speech that is above the listener's hearing threshold or masking noise, whichever is higher; and L_i is the level distortion factor. The formula for calculating L_i is given in Appendix B.

In Annex A of the ANSI S3.5-1997 standard, modifications allow for the fundamental AI to be extended to individuals exhibiting conductive hearing losses. Accounting for this type of hearing loss is accomplished by modifying L_i with a loss factor determined by the conductive hearing loss component of the hearing threshold level.

The relative importance of various frequencies to speech intelligibility for different speech material is provided in Annex B of the ANSI S3.5-1997 standard. The types of speech test material include various nonsense syllables tests (Fletcher and Steinberg, 1929; Pavlovic and Studebaker, 1984; Humes *et al.*, 1986), CID-W22 or PB-words (Studebaker and Sherbecoe, 1991), NU-6 monosyllables (Studebaker *et al.*, 1993), Diagnostic Rhyme Test (Duggirala *et al.*, 1988), short passage of easy reading material (Studebaker *et al.*, 1987), the monosyllables of the Speech in Noise Test (Bell *et al.*, 1992) and average speech (Pavlovic, 1987). These importance functions are shown in Table 3.

In 1996, DePaolis *et al.* attempted to determine the effects of experimental methodology on the frequency-importance functions for differing stimuli. Specifically, they were interested in quantifying the difference between the frequency-importance functions for words, meaningful sentences, and continuous discourse tested under a similar methodology. Recall that frequency-importance functions differ, depending on the stimuli, equipment, and the procedures used to determine them. Twenty-four normal-hearing adults served as subjects and made intelligibility estimates of PB-50 monosyllabic words (ANSI S3.2-1989), sentences from the revised Speech Perception in Noise (SPIN) test (Bilger, 1985), and continuous discourse spoken by a male talker. During data collection, the methodology of low-pass and high-pass filtering at various SNRs closely approximated the procedures reported by French and Steinberg (1947). Results agreed well with other published works,

Table 3. *Band-importance Functions Using Common Audiological Octave and One-half Octave Frequencies for Various Speech Tests: NNS (Various Nonsense-syllable Tests Where Most of the English Phonemes Occur Equally Often), CID-W22 Words, NU-6 Monosyllabic Words, DRT (Diagnostic Rhyme Test), Short Passages of Easy Reading Material, SPIN Monosyllables, and Average Speech. Adopted from Pavlovic (1994) and Reprinted with Permission of Lippincott Williams & Wilkins.*

Band Center Frequency (Hz)	NNS	CID-W22	NU-6	DRT	Short Passage*	SPIN	Average Speech
250	0.0437	0.1549	0.0853	0.0960	0.1004	0.0871	0.0617
500	0.1043	0.1307	0.1586	0.1659	0.2197	0.1190	0.1344
750	0.0841	0.0836	0.0973	0.1142	0.1013	0.1079	0.1035
1000	0.1056	0.1157	0.1068	0.1187	0.0966	0.1078	0.1235
1500	0.1297	0.1349	0.1407	0.1377	0.1092	0.1447	0.1321
2000	0.1664	0.1401	0.1631	0.1288	0.1177	0.1436	0.1328
3000	0.1542	0.1134	0.1244	0.1077	0.1213	0.1433	0.1285
4000	0.1227	0.0648	0.0724	0.0640	0.0676	0.0942	0.1039
6000	0.0893	0.0619	0.0514	0.0670	0.0662	0.0542	0.0796

*This importance function is also known as the importance function for "continuous discourse" (Studebaker *et al.*, 1987) or as the importance function for "easy speech" (Pavlovic, 1987).

indicating that the greatest amount of intelligibility was centered around 2000 Hz for all three stimuli (ANSI S3.5-1969; Schum *et al.,* 1991; Studebaker and Sherbecoe, 1991).

The major finding of the study was that statistically significant differences existed between words and continuous discourse, and between sentences and continuous discourse, when the shape of the frequency-importance functions was compared for one-third octave-band frequencies. This finding suggested that differences in methodology among previous studies were not responsible for different frequency-importance functions, but that the frequency-importance functions were inherently different across stimuli. For octave-band frequencies, frequency-importance functions were not found to be statistically significant for words and continuous discourse. As a result, DePaolis *et al.* (1996) recommended the frequency-importance function for sentences be considered when calculating audibility.

The AI is based on auditory-only communication. Several studies have examined the feasibility of an auditory-visual (AV) AI (AV-AI) (Sumby and Pollack, 1954; Erber, 1969; Binnie, 1973; Binnie *et al.*, 1974; Steele, 1978). For example, Hawkins *et al.* (1988) assessed the perceived intelligibility of conversational speech through the development of a continuous discourse/perceived intelligibility procedure. In the task, normal-hearing listeners provided estimates of percent intelligibility at a number of SNRs. Participants were seated in a moderately reverberant room while continuous discourse was presented from directly in front (0° azimuth) and diffuse noise was presented from 90°, 180°, and 270° azimuths in an auditory-only (A-only) condition. During the A-V condition, a TV monitor displaying the talker was also placed directly in front of the subject. As seen in Figure 11, results showed that speech reading effectively increased SNR by approximately 9 dB. At a -12 dB SNR, the enhancement offered by the A-V over the A-only condition amounts to an approximately 85% improvement in intelligibility performance.

The ANSI S3.5-1969 standard contained a function that relates auditory AI to AV-AI. This function suggested that speech reading adds as much as 0.15 to the A-only calculation of the AI in the lower range. The contribution of the visual modality to speech intelligibility tapers to 0 as the A-only AI approaches 1.0. Grant and Braida (1991) evaluated the adequacy of this function by presenting band-pass filtered speech in quiet and in noise under auditory-only, visual-only, and auditory-visual conditions to a group of normal-hearing listeners. Findings revealed that at an AI score of 0.2, subjects achieved an average of 95% correct in the A-V conditions, but only around 50% correct in the A-only conditions. They concluded that the function in the ANSI S3.5-1969 Standard was adequate only for AI < 0.25.

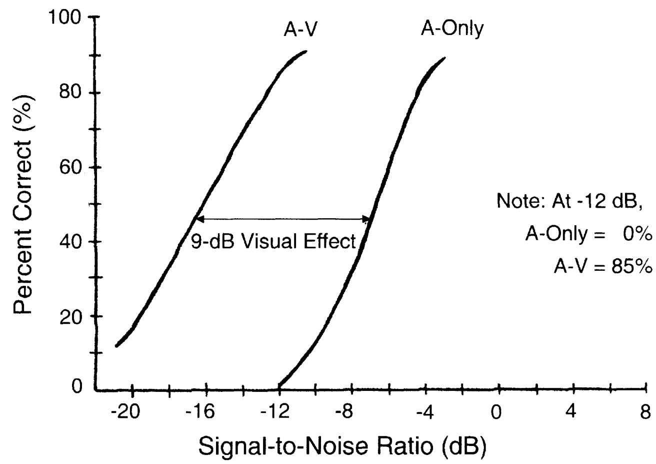

Figure 11. Performance-intensity functions for a listener tested under auditory-visual and auditory-only conditions. Adopted from Hawkins et al. (1988) with permission from first author.

Annex B of the ANSI S3.5-1997 standard also includes an equation for calculating the AV-AI, based on the findings of Grant and Braida (1991):

$$SII_{av} = b + cSII \tag{5}$$

where b and c are fitting constants, depending on the amount of audibility (ie, SII). Figure 12 illustrates this relationship. If SII is $\leq$ 0.2, then b and c are 0.1 and 1.5, respectively. For SII values >0.2, $b = 0.25$ and $c = 0.75$. For example, if one assumes an SII value of 0.35, SII_{av} would result in a value of 0.51 ($b = 0.25$; $cSII = 0.26$ [0.75×0.35]), or an increase of 0.16 with the addition of the visual modality. In general, the contribution of visual cues is assumed to be inversely proportional to the degree of redundancy between the visual and auditory conditions.

In summary, the ANSI S3.5-1997 standard differs from its predecessor in that it accounts for the effects of level distortion, self-speech masking, and upward spreading of masking in the calculation of the audibility function. Furthermore, the ANSI S3.5-1997 standard enables measure-

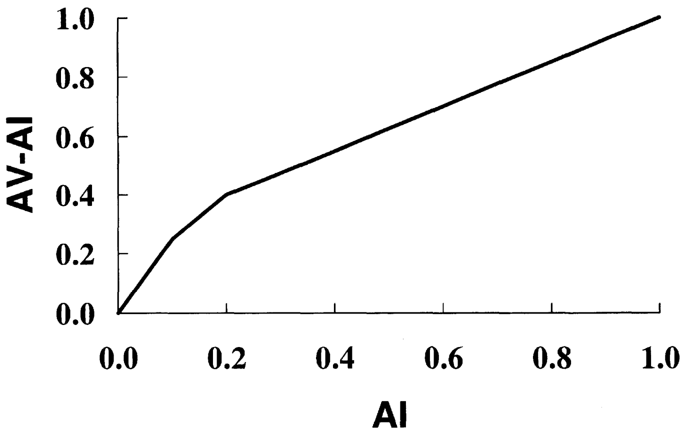

Figure 12. Function relating auditory-visual AI to auditory-only AI.

ments based on modulation transfer functions in reverberant conditions to be used in determining the audibility function. The most significant revision is the provision of various frequency-importance functions for different types of speech test materials. Lastly, a revised function relating the auditory-only SII to an auditory-visual SII is proposed in the newer standard, and it uses a different name to designate the calculated index.

F. Speech Transmission Index and Related Indices

The reader will note that there has been little mention thus far of reverberant environments. Yet, reverberation is similar to competing noise in that it is part of everyday listening environments and it interferes with the intelligibility of target speech. Reverberation time (Rt), or the amount of time required for the sound pressure level of a sound to decrease 60 dB from its offset (ANSI S1.1-1994 [R1999]), has been shown to affect dramatically the speech-recognition capabilities of normal-hearing and hearing-impaired listeners (Moncur and Dirks, 1967; Nabelek and Pickett, 1974; Finitzo-Heiber and Tillman, 1978; Duquesnoy and Plomp, 1980; Nabelek and Robinson, 1982; Neuman and Hochberg, 1983). As seen in the right panel of Figure 13, this interference is created by the reflection of sound energy, particularly low-frequency energy, which causes overlap masking (ie, masking across sounds) and self-masking (ie, smearing of internal energy within a sound) (Nabelek *et al.*, 1989).

To account for the negative effects of reverberation on the AI, correction factors are needed. Because the AI is based on spectral characteristics of undistorted linear speech, it must first be calculated assuming no reverberation. Based on the work of Knudsen (1929), correction factors taking into account the effects of temporal distortion may then be subtracted from the original AI value. These correction factors, which were originally derived under restricted and controlled conditions, have generally yielded poor prediction of intelligibility performance in real-world environments (Humes *et al.*, 1986; Payton *et al.*, 1994). As a consequence, the Speech Transmission Index (STI) was developed as a model for predicting audibility under conditions often found in the real world.

The STI, developed in the 1970s by researchers in the Netherlands (Houtgast and Steeneken, 1971, 1973, 1983, 1985; Houtgast *et al.*, 1980; Steeneken and Houtgast, 1980), is an acoustical index that shares some of the same features as the AI, its predecessor, in addition to having several unique features of its own. The STI calculates SNR using a systems-analysis approach that assesses the relative difference between a well-defined target input signal and the output that is generated within a defined system. In other words, specification of the output relative to the target signal conveys information about the effect of the system under investigation (eg, a room, set of filtering and noise conditions, a hearing-impaired listener, a hearing aid).

The STI uses a target signal that models artificial speech (Houtgast and Steeneken, 1985; Steeneken and Houtgast, 1980). This artificial signal consists of a random noise that has a long-

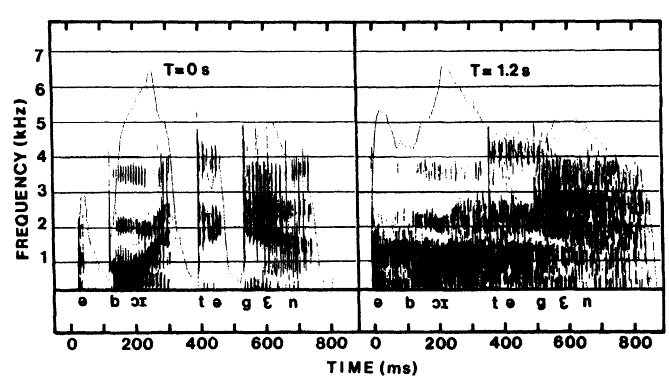

Figure 13. A spectrogram of the phrase "the beet again" without (left panel) and with (right panel) reverberation. Adopted from Nabelek and Nabelek (1994). Reprinted with permission of Allyn and Bacon.

term RMS spectral level similar to that of continuous discourse. Similar to the AI, this target signal is used to measure the SNR in each of seven octave bands centered at 125 to 8000 Hz.

The STI differs from the AI in that the target signal is amplitude modulated, causing temporal fluctuations similar to those observed in everyday communication. Based on the concept that speech can be described as a fundamental waveform that is modulated by low-frequency signals, these temporal modulations incorporate the 14 frequencies of 0.63, 0.80, 1.00, 1.25, 1.60, 2.00, 2.50, 3.15, 4.00, 5.00, 6.30, 8.00, 10.00, and 12.50 Hz (not kHz). The objective of the STI is to determine how much of the modulation is preserved in each octave band and at each modulation frequency at the output.

The amount of modulation preserved is termed the modulated transfer function, or MTF. The MTF can be determined under various conditions such as in a room, under a set of filtering and noise conditions, by a hearing-impaired listener, or through a hearing aid. In essence, the MTF concept is based on the notion that the preservation of envelope fluctuations of the target speech signal is a predictor of speech intelligibility.[3] To determine the effectiveness of the target signal, the amount of modulation of the output signal is compared with the test signal in each octave band. Any reduction in the amount of modulation indicates a loss of intelligibility.

Calculation of the STI begins with determining the MTF for each of the seven octave bands. MTFs are then converted to an average equivalent SNR.[4] Once the SNR is known, the STI is calculated using the equation

$$STI = \sum_{i=1}^{n} I_i \, [(SNR_i + 15)/30] \qquad (6)$$

where I_i and SNR_i are the weighting factors and SNR functions, respectively, for band i. As seen in Equation 6, the STI is similar to the AI (Equation 1) in appearance and in principle (ie, the higher the value, the more audible the signal).

[3]The SII (ANSI S3.5-1997) also incorporates methods of determining speech, noise, and threshold spectrum levels based on MTF measurements. The authors are not aware of any studies that have utilized these methods empirically.

[4]This refers to the specific amount of modulation reduction determined at the output. It is calculated based on the level of the background noise in comparison to the level of the test signal.

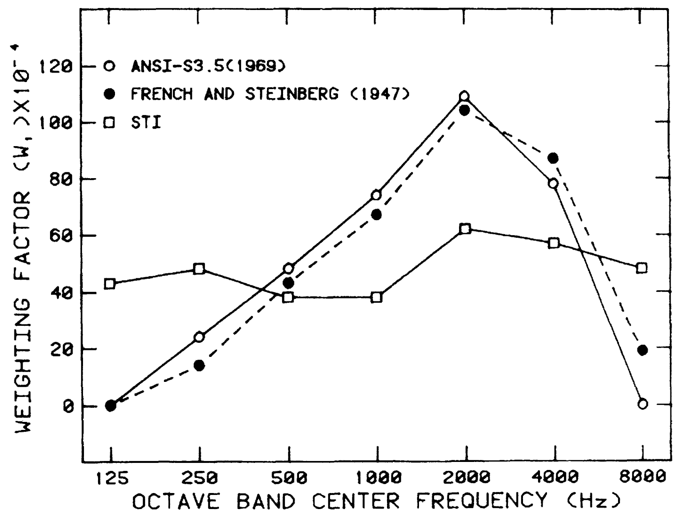

Figure 14. Octave-band weighting factors for monosyllabic words based on two versions of the articulation index (ANSI S3.5-1969; French and Steinberg, 1947) and the Speech Transmission Index (STI). For comparison purposes, the weighting factors have already been divided by 30. Note the near-equal weighting of values for the STI. Adopted from Humes *et al.*, Application of the Articulation Index to the recognition of speech by normal-hearing and hearing-impaired listeners. J Speech Hear Res 29:447-462, 1986. ©American Speech-Language-Hearing Association.

Differences, however, are also evident between the STI and AI. First and foremost, the STI utilizes an SNR of ±15 dB, while the original AI was predicated on a perceptual dynamic range of +12 to -18 dB. (Recall that the ANSI S3.5-1997 standard was recently modified to include a ± 15 dB range.) The STI also differs from the original AI method in its use of frequency-importance weightings, as seen in Figure 14. Specifically, notice that the STI weights are flatter when compared to either the AI (ANSI S3.5-1969) or original weights proposed by French and Steinberg (1947).

The flat spectrum depicted by the STI weightings suggests that each frequency is nearly as important as the other. The 2 AI weightings, on the other hand, ascribe relatively little importance to bands around 250, 500, and 8000 Hz. The difference in weightings between the procedures can be attributed to the fact that the STI was constructed using (a) three similar band-pass filters administered in noise, (b) broad-band speech presented in a broad-band competing noise having a spectral shape that varied from flat to a slope of 12 dB/octave, and (c) several forms of temporally distorted broad-band speech presented against a broad-band noise. Conversely, the AI weights were derived with extreme and abrupt spectral distortion (ie, filtered speech). An example calculation of the STI can be found in Appendix C.

Amplitude compression, such as that found in many nonlinear hearing aids, is also a characteristic whose effects could possibly be assessed by the STI (Humes, 1993). Specifically, amplitude compression decreases the dynamic range of natural acoustic signals into the reduced range between a given hearing-impaired listener's threshold and loudness discomfort level. In a given hearing aid, this compression can be achieved by techniques that filter in the frequency and/or time domains or through digital signal processing (DSP) algorithms that attempt to enhance speech relative to the competing noise.

Experimental findings on amplitude compression have shown no consistent positive effect, and in some cases have shown a negative effect on speech intelligibility, particularly in the presence of reverberation and competing noise (for a review, see Braida *et al.*, 1979 and Plomp, 1988). The STI and MTF are conceptualized as preserving envelope fluctuations of the target signal that contribute to speech intelligibility.

Plomp (1988), for instance, hypothesized that the effect of amplitude compression on speech intelligibility is more detrimental as the number of channels are increased and the compression time constants are decreased (ie, shorter attack/release times). He further argued that the flattening of the envelope fluctuations caused by amplitude compression increases the negative effect of poor frequency resolution on intelligibility. Plomp (1988) suggested that just as the STI/MTF reflects the loss of lower level information due to reverberation and noise, so too it could reflect the loss of spectral and temporal information caused by compression. Consequently, he speculated that the reduced MTF caused by compressing the troughs and peaks of the signal modulation could be used to predict the loss of speech intelligibility due to compression.

In a retort to Plomp's statements, Villchur (1989) indicated that the STI concept does not hold for nonlinear processing. That is, "the weak speech elements that are lost when the MTF is reduced by noise are preserved when the MTF is reduced by compression" (Villchur, 1989, p. 425).

To determine the validity of these viewpoints, attempts to use the STI to predict the performance of amplitude-compressed speech have been undertaken (Steeneken and Houtgast, 1980; Plomp, 1988; Festen *et al.*, 1990; Hohmann and Kollmeier, 1995). In general, findings have demonstrated that the STI fails to describe the effect of amplitude compression on the intelligibility of speech. This shortcoming has been attributed to the stimuli used and the choice of input level, compression threshold, compression ratio, and attack/release times (Humes, 1993; Hohmann and Kollmeier, 1995).

In 1986, Humes *et al.* retrospectively analyzed the results of studies conducted by Moncur and Dirks (1967), Finitzo-Heiber and Tillman (1978), and Duquesnoy and Plomp (1980) to determine the effectiveness of the STI and AI in predicting the speech-intelligibility performance of normal-hearing and hearing-impaired listeners under conditions of temporally and spectrally distorted speech. Findings revealed that the STI provided a better description of speech-intelligibility data under conditions of temporal distortion (reverberation) and conditions not involving abrupt and/or considerable changes in the speech signal. Conversely, the AI was found to provide a superior description of intelligibility for spectrally distorted speech. When the investigators undertook a prospective study that examined the intelligibility of the Nonsense Syllable Test (Resnick *et al.*, 1975) under the conditions of multitalker babble, filtering, and reverberation, their conclusions confirmed the findings of the retrospective analysis.

Based on these results, Humes *et al.* (1986) conjectured that a hybrid model characterizing the best attributes of both acoustical indices might provide a better description of the data obtained in the prospective study. This led to the creation of the modified Speech Transmission Index, or mSTI. Specifically, the mSTI combined the STI's MTF approach of determining SNR with the AI's approach of using 15 one-third octave bands centered at frequencies between 250 and 6300 Hz and the weighting factors originally described by French and Steinberg (1947).

Humes *et al.* (1986) rationalized use of the one-third octave bands by suggesting that octave bands are not as sensitive to the abrupt changes in signal spectrum and that the STI's weighting factors showed no appreciable changes across frequency. In addition to these changes, the mSTI decreased to six the number of modulation frequencies for each one-third octave band. This change also resulted in a wider range of modulation frequencies, which include 0.5, 1.0, 2.0, 4.0, 8.0, 16.0 Hz.

The application of the mSTI was then compared to the results obtained for the STI using data from their normal-hearing group and data acquired from a group of hearing-impaired listeners under quiet and babble conditions. Results revealed that the mSTI offered the best description of speech-intelligibility performance of temporally and spectrally distorted speech in both groups of listeners.

The Rapid Speech Transmission Index (RASTI) is a manufacturer-based derivation of the STI. The RASTI is typically used to predict speech intelligibility in various rooms and other enclosed environments such as auditoriums, theaters, and concert halls. Manufactured by Brüel and Kjaer (1985) as a portable instrument that can make rapid measurements, the RASTI is a simplification of the more complex STI. The simplification is that the RASTI is measured at only two octave bands centered at 500 and 2000 Hz.

Because of this simplification, it has been incorporated as a standard in a number of European architectural codes. Unfortunately, this simplification has forced reevaluation of those standards. For example, because the RASTI tests in only two frequency bands, it assumes that the output response actually extends in a reasonably flat fashion for frequencies lower than 100 Hz and greater than 8000 Hz. While this might be the case in a well-designed auditorium, the performance of many types of paging systems falls short. In these cases, the RASTI almost invariably underestimates the effects of room acoustics on speech intelligibility. Furthermore, compression or limiting in the system can cause artificially low STI and RASTI values by reducing the modulations. This is problematic when, in fact, such processing might enhance intelligibility. It should also be noted that the RASTI does not take system distortion into account.

To date, the RASTI has received little clinical attention because experimentation on hearing-impaired listeners is lacking. Two studies, though, have used this index as a measure of intelligibility in classrooms (Leavitt and Flexer, 1991; Pekkarinen and Viljanen, 1991).

Leavitt and Flexer (1991) measured RASTI values in a moderately reverberant room (Rt = 649 milliseconds). For normal-hearing listeners, they found that the speech-like signal degraded considerably as distance increased from the source. Specifically, a value of 0.83 was determined for the front-row center seat (2.65 meters), 0.66 for the middle-most center seat (6.76 meters), and 0.55 for a student seated 10.88 meters from the source, which equated to the center seat of the last row. In fact, a RASTI value of 1.0 was achieved only at the 6-inch reference position. This finding strongly suggests the need to consider distance when fitting amplification systems.

Pekkarinen and Viljanen (1991) found that RASTI values decreased as background noise and reverberation increased in occupied classrooms. Despite increases in vocal effort by the instructor, speech intelligibility was reported as only fair. The authors suggest use of the RASTI as a mechanism in the design engineering of classrooms.

G. Speech Recognition Sensitivity Model

Recently, Musch and Buus (2001[a], 2001[b]) developed a new intelligibility theory based on a macroscopic model incorporating statistical-decision theory. This theory, called the Speech Recognition Sensitivity (SRS) model, predicts speech-intelligibility performance based on the long-term average speech spectrum, the masking excitation in the listener's ear, the linguistic entropy[5] of speech stimuli, and the number of response alternatives available to the listener. The underlying basis for the creation of this model is that redundant and synergetic interactions among the spectral components of speech account for its intelligibility.

The AI was developed to predict intelligibility performance in telecommunication devices, where each frequency band contributes to the overall signal based on its weighting factor. Each frequency band, therefore, is assumed to contribute independently to phoneme recognition. The SRS, on the other hand, predicts speech intelligibility by combining spectral information from all frequency bands. According to the authors, the SRS model allows for synergetic and redundant information to be simulated, even in the absence of band interactions that may result from the upward spread of masking.

[5]Linguistic entropy is defined as the information content (eg, word frequency, lexical density, ambiguity, sentence "depth", recognition points) of linguistic stimuli (van Rooij and Plomp, 1991).

To accomplish the predictive task, the authors use the concept of statistical-decision theory (Green and Birdsall, 1958), which implies the existence of an *ideal* signal. An ideal signal is best described by means of an analogy. Assuming that a talker and listener share the same linguistic background, the talker will attempt to provide the listener with an ideal signal. Because the accuracy of speech is variable, actual speech usually only approximates this ideal signal. Stated differently, articulated speech typically produces variable degrees of production-related noise, as opposed to ideal speech. This production-related noise, along with any other competing noises, is delivered to the listener's auditory system and determines the intelligibility of the signal. In essence, intelligibility depends on the relation between the intensity of the ideal signal and the intensity of the noises delivered to the ear.

In theory, the SRS model must account for three factors. First, the model assumes that the auditory periphery's ability to decode the incoming signal is correlated with an internal template for every expected word. Second, a linguistic entropy—or knowledge of the phonological, semantic, contextual, and pragmatic variables—is necessary for speech recognition to occur. Third, speech-intelligibility performance is dependent on the number of response alternatives. That is, the smaller the number of alternatives—such as those available in a closed-set listening task—the greater the likelihood that speech intelligibility will increase.

Research has shown that a listener's ability to discriminate among words is related to the type of response format. This relationship is particularly pronounced in closed-set formats consisting of phonemic variations in voicing, place, and manner for each alternative foil (eg, bad, had, sad, mad, pad) and in the number of foils within each test item. Listener performance may be influenced substantially by the response option and number of alternative foils (Studebaker and Sherbecoe, 1993). Thus, predicting scores on materials presented in different formats could result in poorly correlated AI values.

To determine the validity of the SRS, Musch and Buus (2001[b]) assessed the consonant-discrimination ability of 5 normal-hearing listeners. The listener's task was to discriminate among 18 consonants in a CV context, filtered under 58 conditions using either an open-set response format or closed-set response format with 9 foils. Most filter conditions included two or more sharply filtered narrow bands of speech. Using statistical-decision theory, findings revealed that listeners extract speech cues in distinct spectral bands based on redundant, independent, or interactive synergy. This finding was based on the relationship between the observed error rates and the error rates obtained for the individual narrow bands. It was concluded that listeners are more likely to combine speech cues synergistically when a spectral separation occurs between bands and listeners are afforded a closed set of alternatives.

Prediction of Speech Intelligibility for Hearing-Impaired Listeners

A. Hearing Loss and Desensitization

Research has shown the AI to be a valid predictor of speech intelligibility for most listeners with hearing loss in the mild-to-moderate range. Unfortunately, this is not the case for listeners with more severe hearing impairment. A study by Kamm *et al.* (1985) exemplifies these differences.

Kamm *et al.* obtained performance-intensity functions on 5 normal-hearing listeners and 10 listeners with mild-to-moderate sensorineural hearing loss. Each of the listeners had word-recognition scores greater than 88%, as measured clinically in quiet using the NU-6 monosyllabic-word lists (Tillman and Carhart, 1966). An additional listener with moderate hearing loss also participated.

During the experiment, participants were also administered the closed-set Nonsense Syllable Test (NST) (Resnick *et al.*, 1975) processed through a hearing aid capable of simulating both a flat-frequency and high-frequency response at various presentation levels. AIs were computed for each subject at each presentation level and under both frequency responses using the one-third-octave-band method for the NST stimulus items.

Results revealed comparable performance between the observed behavioral measurements and the predicted intelligibility scores for the 5 normal-hearing listeners and the 10 hearing-impaired listeners, each of whom exhibited high word-recognition scores. This finding is in good agreement with results obtained by other researchers (eg, French and Steinberg, 1947; Pavlovic, 1984; Pavlovic and Studebaker, 1984; Humes *et al.*, 1986; Pavlovic *et al.*, 1986; Turner and Robb, 1987; Zurek and Delhorne, 1987; Dubno *et al.*, 1989[b]; Ching *et al.*, 1997).

For the lone moderately hearing-impaired listener with reduced word-recognition scores, the AI was a poor predictor of performance. Similar findings have been observed in listeners with moderate, severe, and profound hearing losses (eg, Dugal *et al.*, 1980; Pavlovic, 1984; Dubno *et al.*, 1989[a]; Ching *et al.*, 1998) and listeners with steeply sloping high-frequency losses (Skinner, 1980; Rankovic, 1991). In general, as hearing impairment increases, the AI overestimates actual speech intelligibility, and the ability to predict AI is overestimated in such cases.

To improve the accuracy of predicting speech intelligibility for these hearing-impaired listeners, various attempts have been made to modify the AI calculation procedure. Fletcher and Galt (1950), for instance, proposed a proficiency factor to account for variations in the enunciation of the talker and in the familiarity of the listener with the talker, using the equation

$$AI = P\, \Sigma I_i A_i \qquad (7)$$

where P refers to the proficiency factor. P reaches its maximum value of 1.0 when communication involves normal-hearing listeners and when the talker and listener speak the same dialect. Because laboratory-recorded materials do not include large variations in talker skills (Studebaker *et al.*, 1995), P is often used as a generalized variable to account for differences between individual listeners that are not explained by audibility (Magnusson, 1996; Magnusson *et al.*, 2001).

Fletcher (1952) proposed in a series of experiments that P should be explored as a means to describe the variability of suprathreshold speech-recognition performance of hearing-impaired listeners, as predicted by the AI. Dugal *et al.* (1980) assessed the efficacy of Fletcher's (1952) proposed proficiency factor by comparing the observed speech scores of 6 hearing-impaired listeners with those predicted on the basis of the AI with and without the proficiency factor. They concluded that the AI that incorporated the proficiency factor improved prediction accuracy for performance at intermediate presentation levels.

Braida (1980), reflecting on the findings of Dugal *et al.* (1980), pointed out that although a constant proficiency factor could accurately analyze data, such a method does not account for the individual degradations in spectral resolution. As a result, Pavlovic (1984) investigated the magnitude of P in accounting for the effects of desensitization by representing the proficiency factor as a function of frequency rather than as a constant value, as originally proposed by Fletcher and Galt (1950). This was accomplished by determining the speech-recognition scores achieved for high-pass and low-pass filtered Harvard Psychoacoustic Laboratory (PAL) PB words (Egan, 1948) by 14 normal-hearing and 16 mild-to-moderately hearing-impaired presbycusic listeners. P was defined as the sum of practice effects and desensitization. In the calculation of P, the practice effect proficiency factor was held constant for both groups, while the desensitization proficiency factor was averaged for subjects' scores on the high-pass and low-pass filtered speech task. Findings revealed that the low-frequency desensitization proficiency factor was not significantly different between listener groups. The high-frequency desensitization proficiency factor, however, was significantly different for the two groups. This finding indicated that desensitization varied as a function of frequency.

To improve prediction based on audibility, Pavlovic *et al.* (1986) suggested a desensitization factor to reduce the deficits in frequency discrimination and temporal coding ability in persons exhibiting cochlear impairment. Basically, this approach, which was derived from earlier work (Pavlovic, 1984), multiplies the frequency-importance function, or I_i, by a correction factor D_i:

$$AI = \Sigma I_i A_i D_i \qquad (8)$$

In the equation, D_i ranges from 0.0 to 1.0 and is derived by means of multiplying the hearing threshold by a factor ranging from 1.0 to 0.0 for hearing thresholds between 15 dB HL and 95 dB HL, respectively. A D_i of 1.0 indicates no desensitization, while a D_i of 0.0 indicates maximal desensitization. The effect of desensitization as a function of hearing loss is illustrated in Figure 15. Using this modified procedure, Pavlovic *et al.* (1986) found substantial improvements, relative to the unmodified AI (1969) method, in the accuracy of absolute predicted values in 4 listeners having various audiometric configurations.

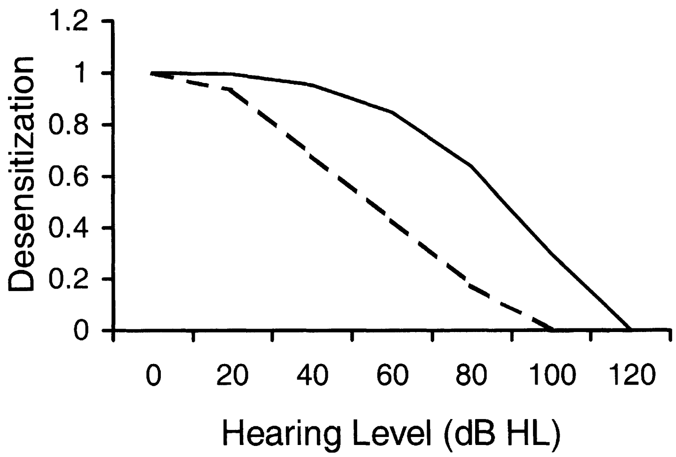

Figure 15. Desensitization as a function of hearing loss, based on the work of Pavlovic *et al.* in 1986 (dashed line) and Studebaker *et al.* in 1997 (solid line).

Ching *et al.* (1997) disagreed with Pavlovic *et al.*'s (1986) notion of a desensitization factor, suggesting that an audible signal contributing zero to intelligibility for hearing losses at and above 95 dB HL seemed inaccurate. They argued that individuals with hearing losses in excess of 94 dB might still be capable of extracting useful information from amplified speech. Ching *et al.* derived a frequency-dependent desensitization factor that could be used to modify the AI for predicting the performance of listeners with differing degrees of hearing impairment.

Eight normal-hearing listeners and 22 listeners with mild-to-profound hearing losses participated as subjects. Subjects were presented with BKB sentence lists (Bench and Doyle, 1979) filtered at three low-pass (700 Hz, 1400 Hz, 2800 Hz) and three band-pass (700 to 1400 Hz, 1400 to 2800 Hz, 1400 to 5600 Hz) speech bands. The speech stimuli were presented at six levels ranging from 6 to 36 dB sensation level (SL) relative to each listener's speech-detection threshold level for each filtered condition. Data collected from normal-hearing listeners were compiled and used to derive a transfer function relating AI to speech scores, and the speech scores of hearing-impaired listeners were compared to the scores predicted by this function. The transfer function was expressed as

$$S = 1 - 10^{(-SII + K)/Q} \quad (9)$$

where *S* is the proportion correct and *K* and *Q* are fitting constants used to minimize the difference between observed and predicted values.

Ching *et al.* (1997) assessed the efficacy of rescaling desensitization as a function of sensation level. Results showed that a smaller desensitization factor could be used to fit behavioral values at high sensation levels or high AIs, and a larger desensitization factor would be required to improve the fitting at low sensation levels or low AIs. The authors found the use of rescaling to be unacceptable because different factors would be required for different sensation levels.

Ching *et al.* (1998) further examined the extent to which speech recognition can be explained by audibility. To do so, they first quantified the speech deficits of hearing-impaired listeners demonstrating a wide range of hearing losses by comparing obtained behavioral performance at various sensation levels to performance predicted by the AI. The sample size consisted of 14 normal-hearing and 40 hearing-impaired subjects. The experimental variables were the same as those described in their earlier investigation (Ching *et al.*, 1997).

As a means to increase the proportion of variance accounted for by the transfer function, the function was further modified, taking the form

$$S = \log_{10} (10^{(k-A*SII)/Q} + 10^{-1/Q})^{-Q} \quad (10)$$

where *S* is the proportion correct, *K* and *Q* are fitting constants to minimize the difference between observed and predicted values, *A* is the slope, and -*Q* is the curvature. Results showed that the modified transfer function provided an accurate description of results from normal-hearing listeners for the various filtered-speech conditions. Some of the hearing-impaired subjects performed better than predicted at low sensation levels or low AIs, while others performed much poorer than predicted at high sensation levels.

As a result of this finding, Ching *et al.* (1998) attempted to address the nonlinear growth of effective sensation level and the discrepancies between observed and predicted intelligibility performance at high sensation levels by proposing several modifications to the ANSI S3.5-1997 standard (Equation 3). Each modified AI procedure was evaluated in terms of the number of fitting parameters required, and the effectiveness of a given formula was determined by calculating the RMS errors between observed and predicted scores of each subject. Ultimately, they recommended use of formulas that improved the relationship between observed and predicted scores by employing the standard level distortion factor (see Appendix B) and individual frequency-dependent proficiency factors. That formula is expressed as

$$SII = \Sigma I_i A_{i,s} L_i P_{n,s} \qquad (11)$$

where L_i is the level distortion factor, or LDF, and $P_{n,s}$ is the proficiency factor for the nth band for subject s.

To evaluate the effectiveness of the unmodified ANSI S3.5-1969 procedure, the ANSI S3.5-1997 procedure, and the modified ANSI S3.5-1997 procedures that incorporate an individual frequency-dependent proficiency factor (Equation 11) for predicting speech-intelligibility performance, the same subjects were tested using BKB sentence lists and VCV syllables that were low-pass filtered at 5600 Hz. The nonsense syllables were comprised of 3 vowels and 24 consonants with 6 repetitions. In the calculation of the SII, the importance functions of sentences and nonsense syllables were used to weight the audibility function (Pavlovic, 1994).

For 17 of the 19 listeners with mild and moderate hearing losses, observed performance was found to be better than predicted at the low SLs of 6, 12, and 18 dB for each of the predicted SII scores. At the higher SLs, observed performance scores were slightly better than the predicted scores for these listeners. For subjects exhibiting severe or profound losses, all three SII procedures, on average, were found to predict poorly the observed performance scores at every sensation level.

Despite a slight improvement achieved by including the effect of a frequency-dependent proficiency factor, results also showed considerable unexplained variance. Overall, these results suggested that audibility, as quantified by the SII, might not be adequate in predicting the speech intelligibility of all hearing-impaired listeners, especially those with severe hearing losses. Amplification that aims to maximize audibility does not optimize speech intelligibility for these listeners.

Studebaker and colleagues (1999) compared the effects of high-level listening on monosyllabic word recognition in a group of 72 young adult normal-hearing listeners, a group of 32 adult hearing-impaired listeners under 70 years of age, and a group of 12 elderly adult hearing-impaired listeners aged 70 years and older. Listeners were presented NU-6 words previously digitized by Studebaker *et al.* (1993) and bandlimited between 447 and 2239 Hz. The stimuli were presented at 8 long-term RMS levels and 10 SNRs relative to a talker-spectrum-matched noise that was band-pass filtered between 282 Hz and 2818 Hz.

Raw data obtained from each subject was transformed into rau units (Studebaker, 1985) and plotted graphically as a function of speech level. Using a method of least-squares regression at each SNR, findings across groups showed that speech intelligibility in noise decreased as the speech level exceeded 69 dB SPL. As a means to predict these outcomes, AI values were derived using an adjusted method. Specifically, SNRs were calculated for each one-third octave band between 125 and 10000 Hz by taking the difference between the speech level and either the applied noise level or the subjects' average internal noise level (ie, threshold).

Fifteen decibels were added to this SNR, resulting in a speech peaks-to-noise ratio (Steeneken and Houtgast, 1980; ANSI S3.5-1997). Each speech peaks-to-noise ratio was then divided by 40, or the effective dynamic range of speech, and multiplied by the frequency-importance function for the NU-6 test (Studebaker *et al.*, 1993). The effective dynamic range of speech value of 40 dB was based on data provided by the normal-hearing group for speech presented in quiet at a level of 69 dB SPL. It was derived by varying repeatedly the assumed range until the mean and standard deviation between the predicted and observed scores showed essentially no variability.

Two optimal estimates of dynamic range were determined (39 and 41 dB) and then averaged. For the normal-hearing group, the AI did not predict accurately observed scores in noise. For the hearing-impaired group, whose data were pooled, results revealed an improvement in the predictability of their observed scores. As a result, it was concluded that the effective dynamic range of speech might be larger than the assumed value of 30 dB.

As a means to understand better the relative importance to speech intelligibility of different intensities within the dynamic range of speech, Studebaker and Sherbecoe (2002) investigated the validity of the 30-dB dynamic range associated with the various AI procedures, as well as the STI. Specifically, they assessed the assumption that the intensity importance assigned to each frequency band is uniformly distributed across the dynamic range and that the dynamic range is the same for all frequency bands.

One hundred normal-hearing listeners were divided into 5 groups of 20 listeners each. Each group provided speech intelligibility data for NU-6 monosyllabic-word lists (Studebaker *et al.*, 1993), sharply filtered into one of 5 frequency bands (141 to 562, 562 to 1122, 1122 to 1778, 1778 to 2818, and 2818 to 8913 Hz). These frequency bands were then mixed with filtered speech-weighted noise and presented alone or in pairs at 19 SNRs ranging from –25 to 47 dB. In essence, each frequency band was presented alone, and also in conjunction with a second band differing in frequency but consisting of the same speech stimuli. For bands presented as pairs, one band was kept at a fixed level, while the other was varied in level. Adjacent bands were never paired.

Data were converted into importance values by determining the amount of change in speech recognition that occurred with and without the inclusion of particular auditory areas of interest. This was accomplished using the equation

$$I = \frac{(A_X - A_Y)}{A_T} \qquad (12)$$

where I is Importance, A_X refers to the speech recognition score obtained when the auditory area of interest was included, A_Y represents the speech recognition score obtained when the auditory area of interest was excluded, and A_T denotes the AI for the entire auditory area.

From this equation, "the importance of each additional dB of SNR within any frequency band was viewed as a proportional part of the total importance of the entire auditory area, not just as a proportional part of an individual frequency band" (Studebaker and Sherbecoe, 2002, p. 1425). It was expected that when speech in two bands added synergistically, the *intensity-importance function* would increase more rapidly with level. In general, results suggested that for the monosyllabic words used in this study, the intensity-importance function varies with frequency and SNR. This view is in contrast to that held presently with regard to the various AI methods, which assumes that intensity is uniformly distributed across each frequency band.

Hogan and Turner (1998) assessed the effect of increasing audibility in high-frequency regions on speech-intelligibility scores for normal-hearing and hearing-impaired listeners exhibiting high-frequency impairments. Fourteen participants, 5 normal hearing and 9 hearing impaired, served as subjects. Each hearing-impaired subject demonstrated no more than a mild hearing loss through 500 Hz. Varying degrees of hearing loss, however, were noted in the higher frequencies.

Subjects listened to nonsense syllables low-pass filtered at cutoff frequencies corresponding to one-third octave bands with center frequencies of 400, 500, 630, 800, 1000, 1250, 1600, 2000, 2500, 3150, and 4000 Hz at 2 different suprathreshold intensity levels (Figure 16). Each subject was presented a minimum of 7 low-pass cutoff frequencies (1120 to 4500 Hz), in addition to a broad-band condition (9000 Hz cutoff). Audibility values were calculated for each by methods in the ANSI S3.5-1969 and ANSI S3.5-1997 standards. Both calculation methods were used to compare the effect of high presentation levels when the level distortion factor was used

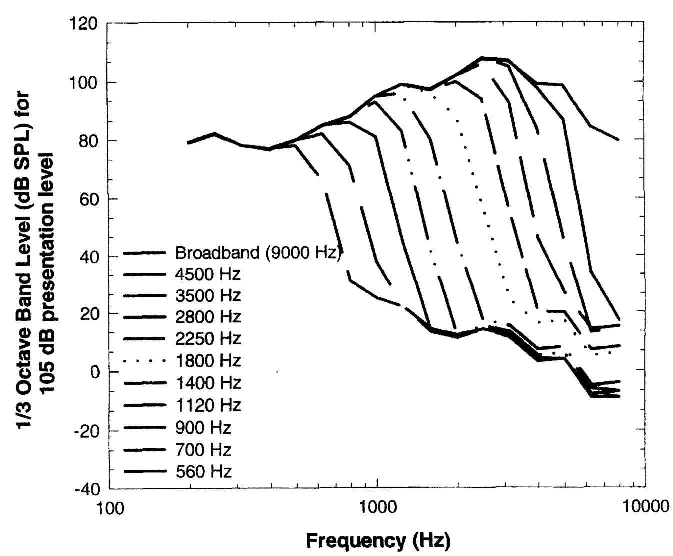

Figure 16. The 12 low-pass cutoff frequencies for the NST spectrum presented at a level of 105 dB SPL. Adopted from Hogan and Turner, High-frequency audibility: Benefits for hearing-impaired listeners, J Acoust Soc Am 104:432-441, 1998. Reprinted with permission of the Acoustical Society of America.

(ANSI S3.5-1997 standard) versus when it was not used (ANSI S3.5-1969 standard). No proficiency or desensitization factors were applied to either method. Speech recognition was quantified, however, using an *efficiency* measure that compared the relative performance of hearing-impaired and normal-hearing listeners as audibility was increased by an additional high-frequency band of speech. Efficiency was defined as "a measure of how well the hearing-impaired listeners used speech information presented at audible levels and at various frequencies compared to the normal-hearing listeners" (Hogan and Turner, 1998, p. 437).

Efficiency was computed using the equation

$$Efficiency = \frac{\Delta\ Score/\ \Delta\ AI\ (hearing\ impaired)}{\Delta\ Score/\ \Delta\ AI\ (normal)} \qquad (13)$$

The formula compares the ratio, in hearing-impaired and normal-hearing listeners, of change in measured intelligibility to the change in AI prediction occurring at higher signal levels. Quantitatively, efficiency was equal to 1.0 if the hearing-impaired listener was able to use the increment in speech audibility from a given band to improve speech intelligibility to the same extent as the normal-hearing group. An efficiency score of <1.0 indicated that the hearing-impaired listener was less able to use the available information to improve the score.

If the increase in bandwidth resulted in rollover (ie, decrease in speech-intelligibility performance), efficiency would result in a negative value. Findings demonstrated an increase in intelligibility scores as audibility increased for normal-hearing listeners and listeners with mild high-frequency hearing losses. They revealed also that listeners with more severe high-frequency hearing losses performed more poorly than normal-hearing, mildly impaired, or moderately impaired listeners.

In addition, comparisons were plotted as a function of the AI methods. Findings from this comparison showed that for a given condition, the newer standard produced lower values, but resulted in more values from the hearing-impaired data falling within the intervals predicted for normal-hearing listeners. This suggests that speech intelligibility can be estimated more adequately from audibility when the effect of level distortion is taken into account.

The authors noted no improvement in intelligibility scores with increase in bandwidth for hearing-impaired participants when hearing sensitivity increased above 55 dB HL. For listeners exhibiting hearing levels below 55 dB HL, information between 4000 and 8000 Hz did not improve speech-intelligibility scores.

Hogan and Turner (1998) concluded that increasing bandwidth may result in a decrease in speech intelligibility for some listeners with high-frequency hearing loss, a finding supported first by the work of Murray and Byrne (1986), and more recently by Ching *et al.* (1998) and Turner and Cummings (1999). In addition, Hogan and Turner (1998) noted a "diminishing return in amplifying high-frequency speech information, in that as amplification is provided in high-frequency regions of moderate and severe impairment, particularly when hearing loss exceeds 55 dB HL, the result may be a decrease in speech recognition performance" (p. 440).

Investigators have attempted to understand the implications of the reduction in speech-intelligibility performance as gain is increased in the high frequencies. Van Tasell (1993), reporting on the work of Liberman and Dodds (1984), pointed out that a decrease in speech-intelligibility performance as hearing loss increased beyond 60 dB HL was evidence of cochlear damage not only to the outer hair cells, but also to the inner hair cells. In fact, it has been conjectured that individuals with similar hearing losses perform differently with hearing aids as a consequence of different degrees of inner hair cell damage (Killion, 1997). Moore (2000, 2001) and Moore and Alcántara (2001) reported that the cochlea of an individual exhibiting a severe hearing impairment could, in fact, have dead regions in which inner hair cells or neurons are not functioning. Unfortunately, such a diagnosis cannot be determined presently based on traditional audiometric test measures.

To circumvent this shortcoming, Moore (2000) developed the Threshold Equalizing Noise (TEN) test. Briefly, this test assesses thresholds for tones in the presence of an ipsilateral broadband noise masker that produces essentially equal masked thresholds, in dB SPL, across a wide frequency range in both normal listeners and listeners with hearing impairment that is unaccompanied by dead regions. A dead region is indicated at frequencies where masked thresholds are 10 dB or more above absolute threshold and 10 dB or more above the equivalent rectangular bandwidth (ERB) level of the masker.

According to Moore (2001), the TEN test is a clinically feasible substitute for psychophysical tuning curves (PTCs), and is useful in isolating regions of the basilar membrane where inner hair cells are not functioning. This test procedure is based on the principle that a cochlear dead region will produce absolute thresholds of the tonal signal when the intensity of the signal is loud enough to produce a detectable spread of excitation at an adjacent region where there are surviving inner hair cells and neurons. Results of the TEN test could shed light on the functioning of the inner hair cells and indicate regions in which amplification is not likely to be useful in improving speech intelligibility. Still in its infancy, the TEN test appears to have potential for future clinical use.

Both the Ching *et al.* (1998) and Hogan and Turner (1998) findings revealed that level distortion and desensitization affect the ability of hearing-impaired listeners to understand speech. Ching *et al.* (2001) addressed the implications of allowing, or failing to allow, for hearing loss desensitization when prescribing hearing aids. They suggested that attempts to maximize audibility at the high frequencies, where the hearing loss is severe, might be inappropriate because the increased gain in regions of severe impairment results in greater loudness, but not necessarily enhanced intelligibility. As a result, clinicians and researchers should aim to maximize *effective audibility,* or the contribution of audibility to speech intelligibility. Effective audibility was incorporated into the SII formula (Equation 3) as follows (Ching *et al.*, 2001):

$$SII_{(LDF,\ HLD)} = \Sigma\, I_i \times \mathit{Effective\ Audibility}_i \quad (14)$$

As seen in this equation, the SII was modified to combine the effects of LDF and a modified desensitization factor, or HLD.[6] These factors were combined based on the findings of previous empirical evidence demonstrating that audibility alone or audibility combined with LDF do not adequately predict the intelligibility performance of listeners with severe hearing impairments. Effective audibility is expressed as

$$\mathit{Effective\ Audibility}_i = \mathit{Desensitized\ Audibility}_i \times L_i \quad (15)$$

Recall that L_i refers to the level distortion factor, LDF. HLD is expressed in terms of desensitized audibility at the sensation level of speech peaks. Thus

$$\mathit{Desensitized\ Audibility}_i = \frac{m_i}{[1+(30/SL_i)^{pi}]^{1/pi}} \quad (16)$$

where SL_i is the difference between the maximum level of the signal and the hearing threshold level at the *i*th frequency band. The factor m_i is the maximum value of desensitized audibility. The rate at which effective audibility changes with audibility at low sensation levels is equal to $m_i/30$. The parameter *pi* is used to control the nonlinear function that relates sensation level to effective audibility. For large values of *pi*, and when $m_i = 1$, effective audibility is equal to the band audibility function as defined in the ANSI S3.5-1997 standard. Both *pi* and m_i vary with hearing loss and frequency, and the relationship between desensitized audibility and audibility for different degrees of hearing losses at different frequencies varies. Figure 17 demonstrates the relationship at 3000 and 4000 Hz.

After deriving effective audibility, Ching *et al.* (2001) compared the observed speech scores of hearing-impaired listeners at each of 6 sensation levels when speech was filtered in a low-frequency band (0 to 700 Hz) and a high-frequency band (1400 to 5600 Hz), and when speech was broadband filtered (0 to 5600 Hz) to predict speech scores based on the original AI (ANSI S3.5-1969), $SII_{(LDF)}$, and $SII_{(LDF,\ HLD)}$ (ie, effective audibility) methods. Findings from the low-frequency band revealed that all methods adequately estimated the performance of listeners with normal hearing and those with mild or moderate flat hearing losses. Performance for individuals with sloping mild or moderate hearing losses was underestimated by both $SII_{(LDF)}$ and $SII_{(LDF,\ HLD)}$ except at the high sensation levels, where effective audibility was found to be a good estimate. For severe and profound flat hearing losses, both procedures that considered LDF adequately estimated performance at

[6]Hearing loss desensitization (HLD) is similar to the term desensitization used throughout this paper. As seen in Figure 15, the difference in terms lies in the fact that Studebaker and Sherbecoe (1994) and Studebaker *et al.* (1997) modified this factor based on speech performance data suggesting that desensitization occurs nonlinearly when hearing loss exceeds 120 dB.

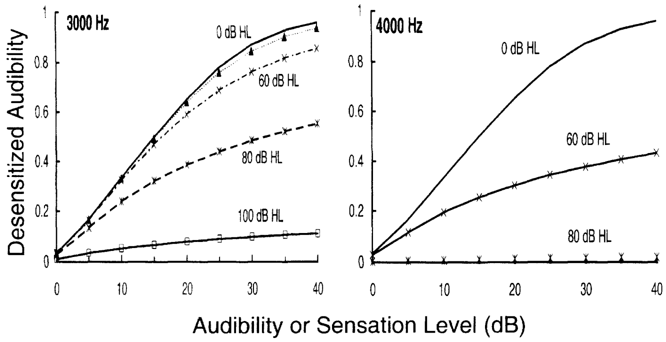

Figure 17. Hearing loss desensitization at 3000 and 4000 Hz as a function of audibility for different degrees of hearing loss (calculated using unpublished data, Ching et al., 2001).

low sensation levels and overestimated performance at high sensation levels. Performance was well predicted by both procedures that considered LDF, but was not well predicted by the AI procedure as specified in the ANSI S3.5-1969 standard.

For the 1400 to 5600 Hz (high-frequency) band, each of the three methods provided adequate descriptions of performance for persons with mild-to-moderate hearing losses. For individuals with severe or profound hearing loss, both the standard AI and $SII_{(LDF)}$ methods were found to overestimate considerably the performance at high sensation levels. For this group, effective audibility was found to be a better predictor of intelligibility performance.

In comparison, for the broad-band speech low-pass filtered at 5600 Hz, the authors found comparable predicted and observed scores for individuals exhibiting a mild or moderate hearing loss. At high sensation levels, effective audibility was found to provide a better estimate of performance for subjects demonstrating severe or profound hearing losses.

These findings contribute significantly to hearing aid fitting applications, particularly as related to the prescriptive fitting of the frequency-gain response and loudness. The clinician must be aware of the implications of providing high-frequency amplification for persons exhibiting severe or profound hearing loss. These recent findings helped to explain why the National Acoustic Laboratories—Revised for profound linear prescriptive method, or NAL-RP (Byrne *et al.*, 1990), which was based on intelligibility judgments and performance of severely hearing-impaired listeners, provided less high-frequency gain than most other prescriptive procedures. The data were used in the derivation of the new NAL prescription for nonlinear hearing aids (NAL-NL1) (Dillon, 1999; Byrne *et al.*, 2001).

B. Age

Another factor to consider in the context of the AI is the age of hearing-impaired listeners. Age is known to affect speech understanding (eg, Plomp and Mimpen, 1979; Dubno *et al.*, 1984; Hargus and Gordon-Salant, 1995; Abel *et al.*, 2000). In fact, some reports suggest that speech-intelligibility performance decreases in listeners as early as age 40, while others have indicated steady performance for listeners through 60 years of age (for a review, see Willot, 1991).

For the elderly population, explanations for decreased speech-intelligibility performance have varied. Overall, this decline in performance can be attributed to a reduction in auditory sensitivity (Humes and Roberts, 1990; Humes *et al.*, 1994), peripheral processing changes (Jerger *et al.*, 1989; Schum *et al.*, 1991; Hargus and Gordon-Salant, 1995), and changes in the central auditory nervous system (Humes *et al.*, 1996). As a result, an age-correction factor might be needed to compensate ". . . for the fact the AI does not predict the performance of older listeners as well as it predicts the performance of younger listeners" (Studebaker *et al.*, 1997, p. 151).

Investigators have assessed differences in speech-intelligibility performance of groups differing in age. Magnusson (1996) investigated the systematic influence of presbycusis and desensitization for 57 elderly subjects, aged 61 to 88 years, exhibiting a sloping mild-to-moderately severe hearing loss. To accomplish this task, the newer AI (ANSI S3.5-1997) procedure was compared to two modified versions of the same procedure using a regression analysis. The first modified version simply took into account desensitization (Equation 8), while the second modified version incorporated desensitization and an age-correction proficiency factor. Age was corrected by the addition of the individual proficiency factor, or P_a. This resulted in the formula of

$$SII = \Sigma I_i A_i D_i P_a \qquad (17)$$

where P_a equates to a proficiency factor of 0.919 for individuals up to 83 years of age. For individuals above 83 years, P_a took the regression form of

$$P_a = -0.0614 \times a + 6.0733 \qquad (18)$$

where a is the age.

The author (Magnusson, 1996) indicates that these age-correction factors suggest no age dependence below 83 years of age and suggest a performance decrease of about 4%, or 2 words in a standard 50-item list, per year for individuals above 83 years of age. The proficiency factor, however, should not be discounted entirely for persons below age 83 years, as it accounts for individual differences not explained by audibility. Magnusson (1996) compared differences between measured and predicted intelligibility scores obtained with the unmodified AI method (Equation 3) and the two age-corrected modified procedures described previously. Results showed a 14% improvement in speech intelligibility for the AI method incorporating the desensitization correction factor when compared to the unmodified AI method. The modified AI incorporating both desensitization and age-correction factors was found to exhibit 21% and 7% improvements, respectively, over the original and initially modified AI procedures. Furthermore, an overall age effect was not apparent for the modified AI incorporating the proficiency factor P_a. This, however, was not the case for individuals above 83 years of age, who demonstrated a decrease in intelligibility performance.

Studebaker *et al.* (1997) also investigated the impact of age on speech recognition by comparing the efficacy of an age-dependent correction factor on the AI. Holding audibility constant, the investigators compared performance across seven age groups (20 to 80 years) using digitized NU-6 monosyllabic words. Audibility was fixed by presenting band-pass-filtered stimuli at a constant SNR and limiting threshold losses to 25 dB HL for frequencies between 250 and 2000 Hz. Eight 50-word lists were presented to each listener on two different days. Raw data were converted to rau scores by using an arcsine transform (Studebaker, 1985). Results revealed that performance did not vary considerably with age, except for subjects over 70 years. For the 70- to 80-year-old group, a modest reduction in scores was noted relative to the 30-year-old group only. A significant reduction in scores was noted for subjects over 80 years of age in comparison to all other age groups. In addition, intersubject variability was found to be greater for subjects over age 40 than for those under age 40. No statistically significant differences attributable to age were noted in learning effects or intrasubject test-retest reliability. Studebaker *et al.* (1997) concluded that the effect of age on the AI, independent of hearing loss, was relatively small.

C. Reverberation

Humes and Roberts (1990) assessed the effectiveness of the mSTI in predicting speech performance in noise for hearing-impaired listeners. They examined the speech-intelligibility performance of young normal-hearing adults, elderly hearing-impaired adults, and young normal-hearing adults with simulated sensorineural hearing loss using the mSTI. Thresholds for the elderly hearing-impaired group consisted of a sloping high-frequency hearing loss.

For the group of subjects tested under the simulated condition, hearing loss was simulated by introducing a spectrally shaped masking noise into the ear receiving the target signal, which resulted in an equivalent loss of audibility seen in the elderly subject group. Subjects listened to nonsense syllables, presented in a closed-set format, recorded in quiet and noise. Speech was presented at 0° azimuth through the left and right ears of an acoustic manikin (KEMAR) (Burkhard and Sachs, 1975), while noise was presented at 0° and 90° azimuth relative to KEMAR.

Recordings were made under anechoic and reverberant (Rt = 3.1 sec) conditions with no noise, noise at 0°, and noise at 90°. Listening conditions for the normal-hearing group and elderly hearing-impaired group included stimulus presentation in the right ear only, left ear only, and binaurally at a presentation level of 75 dB SPL. Data were collected under the left ear only condition for the simulated hearing loss group.

Results revealed that the mSTI, derived using an octave-band method, overestimated the performance of listeners with sloping high-frequency hearing loss, either real or simulated. This was not found to be the case with young normal-hearing listeners in either the anechoic or reverberant conditions. The authors concluded that the use of octave bands and nonsense syllables might have contributed to the overestimation of speech-intelligibility performance.

D. Summary

1. The AI is a valid predictor of speech intelligibility for most listeners with mild-to-moderate hearing impairment. The same is not true for listeners exhibiting more severe degrees of auditory impairment.
2. Severe sensorineural hearing loss results in a reduced ability to extract information that contributes to speech intelligibility, resulting in poor speech-intelligibility performance despite adequate audibility. This reduction is termed desensitization.
3. Modifications of the AI, such as the proficiency factor (*P*) and desensitization factors (Pavlovic *et al.*, 1986, Studebaker and Sherbecoe, 1994; Studebaker *et al.*, 1997), have been derived and assessed as a means to improve predictions of speech intelligibility when desensitization is present. While studies using these modifications have shown improved predicted scores, relative to observed scores, a great deal of variability is noted among hearing-impaired listeners.
4. Findings, in general, have shown that maximizing audibility does not always yield improved intelligibility, particularly in the higher frequencies. In the selection and fitting of hearing aids, a means to overcome this deficiency is through the use of prescriptive formulas that maximize audibility, but yet constrain overall loudness. The NAL-NL1 method was specifically developed with this goal in mind.
5. Regarding age, the average proficiency factor is not much less than 1 for listeners up to age 83 years. Individual differences, however, vary markedly. Above 83 years, the proficiency factor has been found to decrease at a rate of 0.06 (from Equation 18) per year, or a reduction of 4% in the speech-recognition score on a standard 50-item list.

Applications of the Audibility Index to Hearing Aid Selection and Fitting

A. Basic Considerations

Because signal audibility is a fundamental concept in amplification, the AI has been applied clinically to the selection and evaluation of hearing aids. In these applications, Byrne (1992) emphasized the need to standardize signal levels and to compare frequency responses at gain levels typically used by the listener. Figure 18 exemplifies how changes in speech intensity level produce changes in the AI, using the A_0 method devised by Pavlovic (1988). In the left panel, an individual's hearing thresholds are plotted in relation to an overall speech level of 63 dB, giving an AI of 0.17. When the same hearing thresholds are plotted in relation to an input level of 70 dB, the AI increases to 0.38, as shown in the right panel. Relatively small changes in signal level, whether based on speech input level, or in the case of aided listening, on the gain of a hearing aid, can produce substantial changes in the AI. Which speech levels best approximate those encountered by hearing-impaired listeners?

Byrne and Cotton (1987) found that as hearing sensitivity increased for hearing-impaired listeners, an increase in speech input level was needed. Specifically, they found that aided listeners with a 3-frequency (500, 1000, and 2000 Hz) pure-tone average of 100 dB preferred an increase in speech input level that differed by only 7.5 dB when compared to aided hearing-impaired listeners with mild-to-moderate hearing loss. Assuming that the speech presentation level typically received by a person with normal hearing is

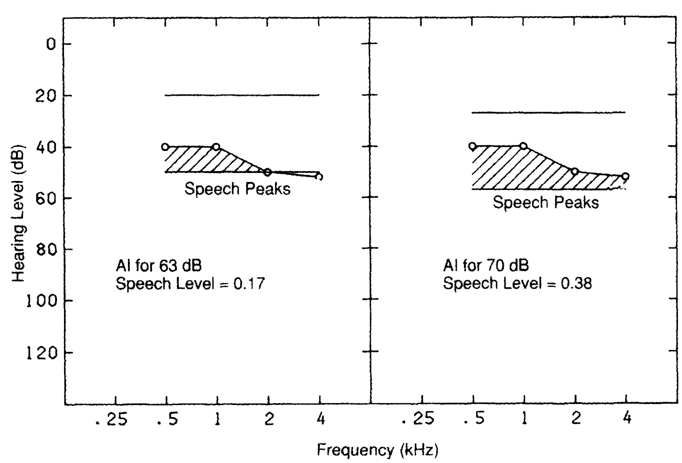

Figure 18. Illustration of the $A_0(4)$ procedure (Pavlovic, 1988) using the normal range of speech (left panel) and a higher range of speech (right panel). Adapted from Byrne (1992). Reprinted with permission of the *Journal of the American Academy of Audiology.*

65 dB SPL, the presentation level received increases to 69.5 dB for a 60 dB hearing loss, and to 71 dB for an 80 dB hearing loss.

Based on the findings of Byrne and Cotton (1987), intensity levels of 65 dB to 70 dB and levels of 70 dB to 75 dB appear to be adequate for mild-to-moderate hearing losses and severe hearing losses, respectively.

AI will also change as the frequency response, or gain-by-frequency, of a hearing aid varies. In Figure 19, an aided threshold curve is depicted within the 30-dB dynamic range for a given speech stimulus. Note that the aided AI is determined to be 0.73. Any changes to the gain or frequency response, however, will also result in changes to the AI. As an example, assume that the listener enters a different acoustical environment requiring increased gain. Assume further that the listener adjusts the volume control, which results in a 3-dB increase at each frequency. This change in frequency-gain response produces a relative aided AI increase of 0.10, or an overall AI of 0.83. Suppose the listener now leaves this acoustical environment and enters a new one. The newer environment is perceived as being too loud and the volume control is manipulated again, this time to decrease the entire frequency-gain response by 6 dB. This 6-dB decrease in amplification results in a 0.20 reduction in the AI. Therefore, clinicians must be aware that prescriptive formulas, whether used in the selection or the fitting of a hearing aid, attempt to provide optimal gain and, inherently, the highest AI value.

Reliance on the prescriptive method or an assumption that the patient will persist in adjusting the volume control to maximize AI is irrational. Byrne and Cotton (1987) point out that data on preferred listening levels suggest that the AI will rarely, if ever, be maximized for listeners with greater than a moderate hearing loss. In fact, they found that individuals with mild hearing loss did not always maximize AI for various listening conditions.

The manner in which the aided frequency-gain response is viewed, either relative or absolute, can also lead to incorrect conclusions about the AI. Absolute measures refer to those made routinely on hearing aids, whereas relative measures refer to comparisons of differences between frequency responses of two or more hearing aids. In hearing aid selection and fitting, Byrne (1992) noted that relative measures reduce

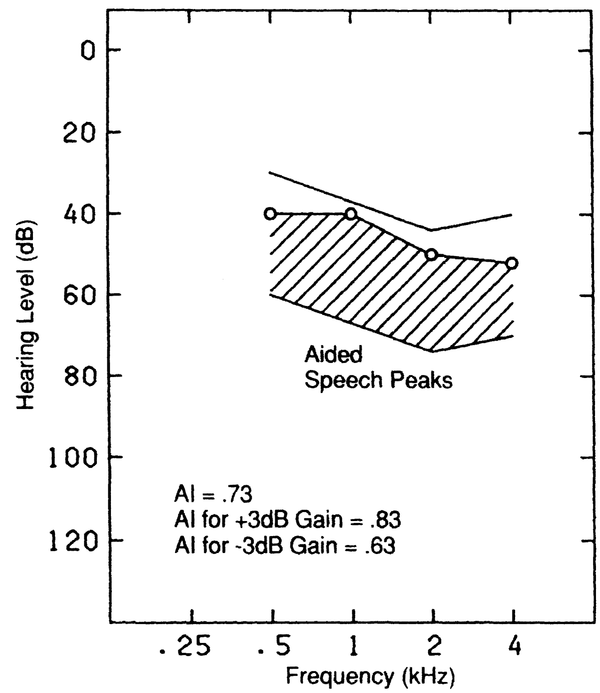

Figure 19. Calculation of AI for aided listening. Adapted from Byrne (1992). Reprinted with permission of the *Journal of the American Academy of Audiology.*

procedural complications, but that incorrect conclusions can result from both types of measures.

Using relative measures as an example, Figure 20 depicts an SPL-O-Gram with threshold and discomfort levels as the minimum and maximum, respectively, and speech-peak levels for 2 frequency responses after speech has been adjusted to the preferred level for each response. Using Pavlovic's A_1 (1989) method, a computerized procedure that uses an importance function for continuous discourse, notice that frequency response 1 (solid line) allows for greater low-frequency and high-frequency gain compared to frequency response 2 (dashed line). Assume that response 1 is the *target* level and response 2 is the listener's *preferred* level. Because the midfrequencies are important predominately for speech intelligibility, a first impression might be that response 2 is better. However, by comparing the relative overall difference between frequency-gain response curves, both responses in this example produced AI values of 0.36.

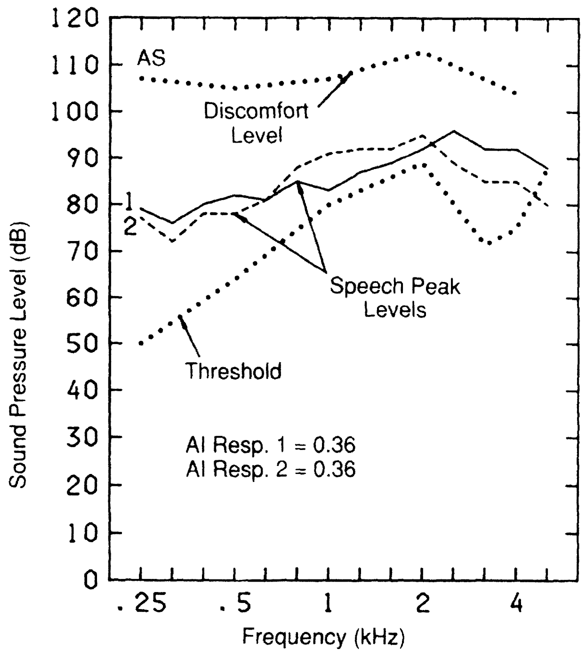

Figure 20. Example of speech-peak levels obtained with two frequency responses, each adjusted to the preferred listening level. The listener's threshold and discomfort levels are also shown. Adapted from Byrne (1992). Reprinted with permission of the *Journal of the American Academy of Audiology.*

Similarly, incorrect conclusions can also be drawn when considering only the absolute values. As seen in Figure 21, AI values have been calculated for the same two frequency-gain responses seen in Figure 20, but at a nonpreferred gain of 6 dB lower. Here, the AI value is lower for response 2 (0.17) than for response 1 (0.23). When compared to Figure 20, also notice that the AI value for response 2 is more reduced (difference of 0.19) than the AI value for response 1 (difference of 0.13). Based on the AI values, we would conclude that response 1 is superior to response 2. By incorporating a relative comparison, one can clearly see that response 1 would provide no amplification in the critical mid-frequency region, while response 2 would provide some amplification in that region, though output levels would be near threshold.

From these examples, the attractiveness of using the AI in hearing aid work is clearly evident because it describes the amount of speech information (ie, speech spectrum) available to the listener (Pavlovic, 1988). However, it should be noted that this feat has been accomplished using the original AI (ANSI S3.5-1969) formula in Equation 1. Recall that in this equation, the proficiency factor and any other derivatives are absent. That is, correction factors for variables such as level effects or the upward spread of masking are not usually incorporated. As a result, increases in audibility as a result of amplification may not always provide an improvement in speech-intelligibility performance.

B. Evaluation of Frequency-Gain Response

In the clinical selection and fitting of hearing aids, investigators have typically attempted to use the AI in one of two ways: (a) as a basis for selecting one hearing aid frequency-gain response over another, or (b) as a model that will specify a frequency-gain characteristic that will optimize speech-intelligibility scores.

Studies have compared hearing aid fitting prescriptions through the use of the AI model in the evaluation of optimal frequency-gain charac-

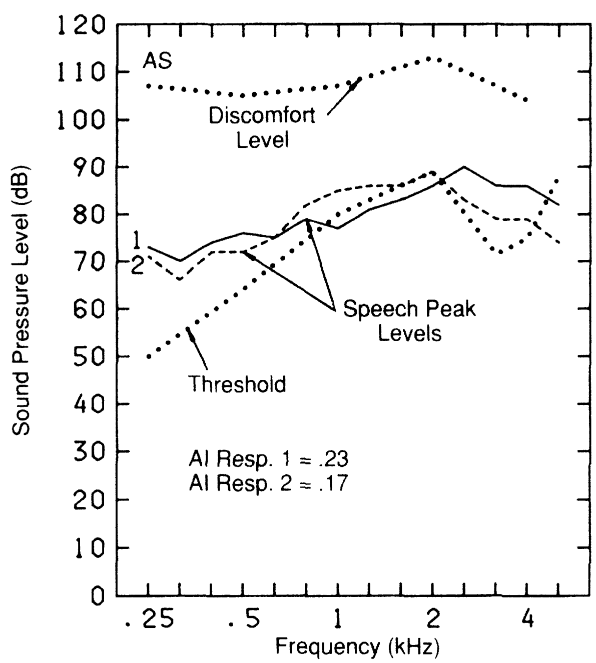

Figure 21. Variation of Figure 20. Note that the speech peak levels have been reduced by 6 dB for both responses. Adapted from Byrne (1992). Reprinted with permission of the *Journal of the American Academy of Audiology.*

teristics. Although these investigations address indirectly the intelligibility of speech, their direct focus is on audibility. Humes (1986), for example, compared prescribed insertion-gain values derived from 10 linear hearing aid selection procedures as applied to 9 different degrees of audiometric configurations. The hypothetical audiograms for each of the 9 cases were grouped by their slope configurations of sloping, flat, and rising, resulting in 90 sets of prescribed insertion-gain data. These insertion-gain data were then compared to determine how well each procedure achieved the goal of maximizing speech intelligibility. The results of this analysis showed that when volume-control adjustments were not allowed, 8 out of the 10 prescriptive methods prescribed sufficient gain to yield optimal speech-intelligibility performance. The insertion-gain data were then compared to hearing-impaired listener preferences based on the work of Leijon *et al.* (1984). Interestingly, findings revealed that the same two procedures found not to provide optimal speech-intelligibility performance were most preferred by these listeners. It was also surmised that when the listener is given the ability to adjust the overall gain with the use of a volume control, any of the procedures could produce optimal aided speech-intelligibility performance.

In 1991, Rankovic compared the frequency-gain response of two linear prescriptive procedures and a prescription that aimed to maximize the AI (AIMax). The hearing aid prescriptive procedures were the Prescription of Gain/Output (POGO) (McCandless and Lyregaard, 1983) and the National Acoustic Laboratories–Revised procedure (NAL-R) (Byrne and Dillon, 1986). POGO is essentially one of a handful of half-gain methods, while NAL-R was derived from a considerable amount of clinical and laboratory research aimed at amplifying the long-term average speech spectrum to a comfortable equal-loudness contour, predicted from pure-tone thresholds (Byrne, 1986[a], 1986[b]).

AIMax was aimed at maximizing the frequency-gain response by positioning the 30-dB dynamic range of the short-term speech levels 15 to 18 dB above the listener's pure-tone thresholds and below the level of discomfort. This was achieved by selectively amplifying the speech spectrum so that the long-term average one-third octave-band level of speech in each band was 18 dB above the pure-tone threshold and below a formulated UCL.

As a result, subjects were expected to receive "the minimum amount of gain that assured maximum audibility of the speech spectrum in each frequency region" (Rankovic, 1991, p. 392), in principal, an AI of 1.0. Twelve hearing-impaired subjects, each exhibiting various degrees and configurations of hearing loss, participated in the study. No subject achieved an AI of 1.0 because of the noise floor, limitations of the dynamic range of the equipment, and/or subjects' UCLs. The relationship between the percentage of correct scores on a nonsense syllable test and audibility was compared for all subjects. Overall, findings showed AIMax to be similar to or slightly better than POGO and NAL-R, possibly a result of its requiring more prescribed gain than the other prescriptive methods. AIMax appeared to provide adequate gain for individuals exhibiting relatively flat hearing losses, although data did not reveal if these listeners found the AIMax condition to be comfortable during everyday listening conditions. While this increase did not always yield an improvement in observed scores, it also did not degrade performance for most listeners. Four of the subjects, each demonstrating a sloping high-frequency hearing loss with near-normal low-frequency thresholds, demonstrated a decrease in speech-intelligibility performance for AIMax. This decrease in performance occurred as a result of the considerable amount of gain needed in the higher frequencies to achieve maximum intelligibility. Skinner (1988) also noted a similar finding for persons with high-frequency hearing loss.

Linear prescriptive approaches fail to address the fact that changes in speech spectra, ambient noise, and other environmental factors are commonplace in everyday listening situations. Hou and Thornton (1994) developed a method for integrating the AI as a means to predict speech intelligibility across a range of listening conditions. This integrated AI model, termed IAI, takes into account hearing threshold, masking of noise, self-masking of speech, high-level cochlear distortion, and peak-clipping effects of a hearing aid. The IAI evaluates separately AI values of hearing aid performance across several listening conditions. These derived AI values are first grouped according to such factors as listening conditions and input levels, and then weighted based on importance. The sum of these groupings and weightings is then integrated to provide an estimate of the optimal performance, or optimal IAI (OIAI), for a specific hearing aid or hearing aid characteristic.

Hou and Thornton (1994) validated their approach by using the integrated AI across a range of listening conditions as a criterion for evaluating a specific hearing aid response characteristic and calculating an optimal frequency-gain characteristic that maximizes the IAI. For a high-frequency hearing loss, the OIAI was compared to the NAL-R (Byrne and Dillon, 1986) and the POGO (McCandless and Lyregaard, 1983) prescriptions. The results in quiet showed no difference among the three procedures in predicted performance, but in a noisy condition, superior performance was predicted for the OIAI procedure when compared to the two prescriptive procedures.

As a result of the examples presented in this section, the reader should readily appreciate that the selection and fitting of the electroacoustic characteristics of hearing aids should be carefully considered. Simply stated, the highest AI value may not provide the listener with optimal speech intelligibility. Empirically, Byrne (1986[b]) found this to be true. He assessed judgments of quality and intelligibility in quiet and noise for 14 ears of 11 participants, 7 of whom exhibited precipitous hearing losses. Frequency responses yielding greater high-frequency emphasis produced the highest AI values. However, six of the seven listeners with precipitous loss judged the frequency responses yielding the highest AI value to be the poorest in three of four trials based on paired-comparison preference judgments. In 1991, Rankovic confirmed this finding when participants with high-frequency hearing losses did not achieve the best speech recognition with a frequency response that maximized AI.

C. Prescription of Frequency-Gain Response

Researchers have also attempted to use the AI model to specify a frequency-gain characteristic for optimizing speech-intelligibility scores. Humes (1986) and Rankovic (1991) have both reported the use of procedures that aimed to amplify the full 30-dB range of speech to be audible at each frequency for hearing-impaired listeners. These procedures have not been adopted for clinical use.

At present, the only prescriptive procedure that uses an AI model for deriving optimal frequency-gain characteristics is the NAL-NL1 procedure for fitting nonlinear hearing aids (Dillon, 1999). This prescription aims to maximize speech intelligibility using a modified AI procedure (ANSI S3.5-1997) that incorporates level distortion and hearing loss desensitization (as specified in Ching *et al.*, 2001), within the constraint of an overall loudness that is either the same or lower than that experienced by a normal-hearing listener at the same input level (Dillon, 1999; Byrne *et al.*, 2001). This prescriptive procedure can be implemented via standalone software or via some manufacturers' fitting software. Like any other such procedure, the prescription is based on group data, and does not account for individual differences. It should be viewed, therefore, as providing a starting point, with fine-tuning performed as needed for individual hearing aid users. Because NAL-NL1 is still in its infancy, its effectiveness in predicting the speech-intelligibility performance of aided listeners has not yet been validated fully.

D. Evaluation of Amplification Strategies

In 1990, Fabry and Van Tasell used the AI to evaluate a hearing aid incorporating adaptive frequency response (AFR). AFR is a hearing aid circuit that can be enabled to reduce automatically the low-frequency gain in noisy backgrounds. The authors chose to investigate this feature due to its widespread application throughout the industry at the time, despite their acknowledgement that articulation theory generally has not supported this type of signal processing as an effective noise-reduction technique (French and Steinberg, 1947; Fletcher, 1952). Using connected discourse, a transfer function relating rated speech intelligibility to the AI was derived for 12 normal-hearing subjects. This transfer function was then used to predict aided speech-intelligibility ratings by 12 hearing-impaired listeners wearing a master hearing aid with an adaptive filter circuit that could be activated (AFR on) or deactivated (AFR off). For all subjects, results showed that the AI predicted no improvements in speech-intelligibility performance for the AFR-on versus AFR-off condition. In addition, no significant improvements in rated intelligibility were observed. For every hearing-impaired participant, however, ratings of speech intelligibility were related monotonically to AI.

E. Prediction of Hearing Aid Outcomes—Speech Gain

In an attempt to overcome the need for deriving proficiency factors and the various transfer functions associated with speech stimuli, Dillon (1993) formulated and evaluated a means by which *speech gain* provided by a hearing aid can be predicted from insertion gain. Speech gain was defined as "the difference in level between the aided and unaided performance-intensity (PI) functions measured at any specific value of percentage of items correct" (Dillon, 1993, p. 621). Using 11 listeners with mild or moderate sensorineural hearing loss, AI was used to predict gain based on unaided sound-field thresholds, ambient room noise, hearing aid internal noise, measured real-ear insertion gain, and unaided PI curves. Unaided sound-field thresholds were obtained in dB SPL at octave intervals from 125 to 500 Hz and at one-half octave intervals from 750 to 6000 Hz. Insertion gain was measured at one-third octave frequencies from 125 to 8000 Hz.

Every participant was tested with monosyllabic iso-phonemic AB lists (Boothroyd, 1968), 9 of the subjects received the continuous discourse stimulus read by a male speaker, and only 1 subject was given the 36 nonsense syllable vowel-consonant-vowel items spoken by a female talker. During the testing procedure, unaided and aided PI functions were observed for each stimulus at percent-correct values of 25%, 50%, and 75%, and the corresponding points plotted. Speech gain was determined simply as the difference, in dB, between the measured unaided and measured aided PI curves, as illustrated in Figure 22.

Dillon (1993) calculated unaided AI by considering the minimum audible level in each one-third octave band as the maximum measured room noise, equivalent input noise of the hearing aid, and the equivalent internal noise corresponding to the hearing thresholds. Sensation level within each band was defined as the maximum speech level minus the minimum audible level over a 30-dB range. Importance functions believed to be applicable to nonsense syllables and continuous discourse were used to derive predicted scores. Findings showed that measured speech gain could accurately be predicted from electroacoustic information consisting of the subject's thresholds, ambient noise in the test environment, internal noise of the hearing aid, and the device's insertion gain. Predicted scores differed from observed scores by an RMS error of 3 dB—or about 6 dB for 5% of patients—at each of the three hearing levels for any of the speech stimuli. The choice of importance function was not critical to the outcome.

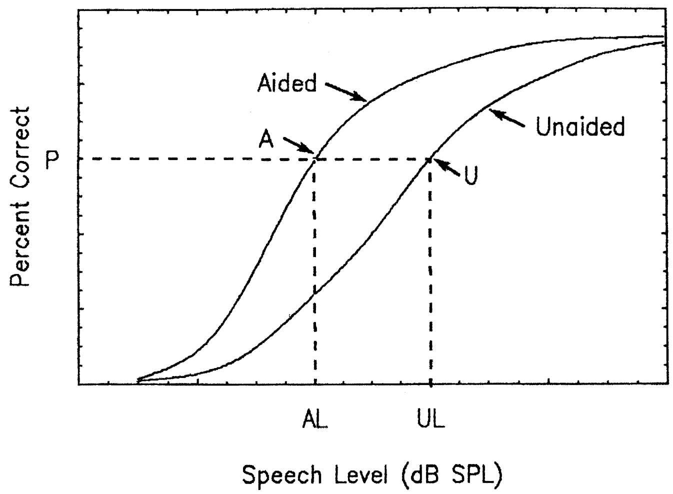

Figure 22. Illustration of the aided (AL) and unaided (UL) levels for predicting speech gain. Adopted from Dillon, Hearing aid evaluation: Predicting speech gain from insertion gain, J Speech Hear Res 36:621-633, 1993. ©American Speech-Language-Hearing Association. Reprinted with permission.

Additional advantages of using this predicted speech gain method over the traditional method of predicting speech recognition from the AI are that REIG measures are readily obtainable, and perhaps more importantly, substantial time can be saved by determining the speech gain provided by a hearing aid from electroacoustic measures. Dillon also pointed out that those listeners with greater hearing aid experience attained higher speech gains than predicted. Listeners with less than 2 months of experience, on the other hand, tended to exhibit lower than predicted scores. A considerable limitation to this method, however, is that "Insertion gain and narrow-band noise hearing thresholds cannot be used to predict actual speech comprehension and/or remaining hearing deficit" (Dillon, 1993, p. 630).

The AI is limited in its ability to account for compression-amplified speech. This is due simply to the fact that amplitude compression reduces the dynamic range of speech as intensity increases, thus violating the 30-dB assumption on which the AI is based. In recent years, the AI has also been used to determine which nonlinear prescriptive procedure best optimizes audibility through the selection of one hearing aid frequency-gain response over another. Most such studies have assessed the audibility provided by wide dynamic range compression (WDRC) circuitry. These circuits provide relatively more gain for low-level input signals than for high-level input signals. This feature, in theory, provides the listener with less of a need to adjust the volume control. Moore (1996) conjectured that in WDRC systems with sufficiently fast attack and release times, the compressor reduces the intensity differences between low-level and high-level components of speech, which in turn, may provide the listener with improved audibility of low-intensity phonemes (ie, consonants).

Stelmachowicz *et al.* (1998) noted that when fitting WDRC devices, clinicians typically gather loudness judgments to set the input/output characteristics of the hearing aid circuit. Unfortunately, some children and elderly individuals do not possess the cognitive skills needed to make these judgments. To determine the best way to predict the appropriate frequency-gain characteristics for these individuals in quiet, Stelmachowicz and colleagues analyzed the fittings for 49 adult hearing aid users with mild-to-severe cochlear hearing loss. Loudness-growth measures were obtained for every subject using the loudness growth in octave band (LGOB) algorithm (Allen *et al.*, 1990; Pluvinage, 1994). Frequency-gain characteristics at 50 and 80 dB SPL were derived based on a manufacturer's threshold-based fitting scheme and two nonlinear prescriptive procedures: Fig6 (Killion and Fikret-Pasa, 1993) and DSL[i/o] (Cornelisse *et al.*, 1994, 1995). Results showed that both Fig6 and DSL[i/o], in general, provided more gain than was actually used by the listeners, and discrepancies increased as a function of frequency. Furthermore, it was shown that the LGOB algorithm provided a closer approximation to the amount of use gain than either nonlinear prescriptive method.

Using these prescriptive gain values, Stelmachowicz *et al.* (1998) assessed audibility through the use of the Aided Audibility Index (AAI) for soft (50 dB SPL), average (60 dB SPL), and loud (70 dB SPL) speech across various degrees of hearing loss. The AAI for WDRC[7] takes the form

$$AAI = [\sum_{i=1}^{8} I_i\,(SL - LIMIT - DF)]/RANGE \qquad (19)$$

In this equation, I represents the band-importance values and i equates to the 6 one-third octave band frequencies and 2 interoctave band frequencies (3000 and 6000 Hz) commonly used audiometrically. *SL*, or the amount of audible speech in each band, is obtained by subtracting the listener's one-third octave-band threshold from the speech maxima in that band. The speech maxima, as defined by this model, have no fixed values and are based on the amount of compression. That is, the higher the compression ratio, the more reduction in speech peaks. *LIMIT* accounts for the OSPL90 value of the hearing aid and is calculated by subtracting the OSPL90 from the speech peaks in that band. *DF* is the distortion factor due to peak clipping, based on the work of Crain and Van Tasell (1994). Finally, *RANGE* refers to the speech range, which is less than 30 dB when compression is used.

Stelmachowicz and colleagues (1998) used the transfer functions for average conversational speech and nonsense syllables (ANSI S3.5-1969) to predict performance for persons with mild-to-moderately severe hearing losses as a function of input level. For these hearing losses, they found DSL[i/o] to provide a more consistently audible signal than the other two threshold-based algorithms.

Using the already described AAI in a retrospective analysis of earlier-published data, Souza and Turner (1998) examined the relationship between audibility and speech-intelligibility performance for compression-amplified versus linear-amplified speech. The motive for their undertaking was based on findings in the literature indicating that under some conditions, amplitude compression increased the amount of audible information over that of a linear amplifier, but did not always result in improved recognition scores. Souza and Turner (1998) speculated that these findings might be attributable to the alterations of the speech envelope at high intensity levels relative to the input signal. As a result, relative improvements in recognition based on increased audibility for compression-amplified speech would not be expected when compared to linear-amplified speech.

For the linear-amplified condition, the highest input level was amplified to a 100 dB SPL output

[7]An AAI for linearly amplified speech also exists (Stelmachowicz *et al.*, 1994). The difference in this formula, relative to the AAI for WDRC, is that the *SL* factor's speech bands equal the long-term average speech spectrum +15 dB. In addition, the *RANGE* value is equal to 30, or the presumed dynamic range of speech.

level. During pilot testing, Souza and Turner (1999) found this level to provide maximum audibility without discomfort for listeners with similar degrees of hearing loss. As a result, the *LIMIT* and *DF* factors were omitted for both AAI calculations (Equation 19). A two-channel compressor with a 1500 Hz center frequency was used for the compression-amplified condition. Each channel of the WDRC device compressed signals above 45 dB SPL and at compression ratios of 2:1 and 5:1 in the low-frequency and high-frequency channels, respectively.

Sixteen hearing-impaired subjects heard 16 VCV stimuli spoken by 2 male and 2 female talkers at presentation levels of 55, 70, and 85 dB SPL for both amplified conditions. Each subject's task was to indicate recognition of the correct consonant in a 16-alternative forced-choice paradigm. Using the appropriate AAI measurement, the relationship between audibility and intelligibility was examined for each type of amplification.

Souza and Turner (1999) found that at low and moderate input levels, AAI and intelligibility scores were higher for compression-amplified speech relative to linear-amplified speech. At the higher input level, however, AAIs and intelligibility scores were similar for the two types of amplification. Because intelligibility scores increased monotonically for both types of amplification, no significant differences were found between compression-amplified speech and linear-amplified speech. This finding suggests that a given increase in audibility resulted in the same increase in intelligibility for both amplification conditions. Clinically, this outcome offers preliminary evidence that compression does not introduce detrimental changes to the speech signal that might otherwise reduce audibility.

Recently, Souza and Bishop (2000) took the work of Souza and Turner (1999) one step further. Specifically, they attempted to determine whether increases in audibility with nonlinear amplification improved speech intelligibility to a comparable degree for 10 listeners with sloping sensorineural hearing loss relative to a group of 10 listeners with a flat sensorineural configuration. Except for the selection of a frequency-gain response that maximized high-frequency audibility, the methods and AAI calculations were similar to those used in the Souza and Turner (1999) study.

For linear-amplified speech, overall results revealed that listeners with flat and sloping hearing losses showed similar improvements in intelligibility, given comparable increases in audibility. At the high-intensity level, however, listeners with high-frequency hearing loss showed smaller improvements in intelligibility with nonlinear amplification when compared to listeners exhibiting flat audiometric configurations.

The authors offered two possible explanations for the difference between groups. First, empirical evidence has shown that listeners with sloping hearing losses have a compromised ability to use temporal cues at high intensity levels (Bacon and Viemeister, 1985; Bacon and Gleitman, 1992). Second, recalling that a greater compression ratio of 5:1 was used in the high-frequency channel, time-amplitude cues in this region may have been affected the most, which in turn, might have led to a further reduction in audibility.

In 2001, Souza and Kitch compared preferred volume-control settings for (1) a linear peak-clipping, (2) compression-limiting, and (3) WDRC multimemory, programmable behind-the-ear (BTE) device. For each amplification type, speech audibility was quantified at the listener's preferred volume setting. Ten listeners with mild-to-moderate cochlear hearing loss were fitted monaurally. The frequency response, OSPL90, and WDRC ratio were set individually based on the DSL [i/o] (Cornelisse *et al.*, 1995) prescriptive procedure.

For the compression-limiting hearing aid condition, the aid was set to an output-compression mode, and compression thresholds were programmed between 65 and 85 dB SPL, with a compression ratio ranging between 8:1 and 20:1 relative to the listener-preferred volume-control setting. An input-compression mode was used for the WDRC condition, with the compression threshold fixed at 50 dB SPL and compression ratios ranging from 1:1 to 2.7:1.

Subjects listened to 20 passages of the Speech Intelligibility Rating (SIR) test (Cox and McDaniel, 1989) presented at input levels of 50, 65, and 80 dB SPL in quiet and against the SIR's accompanying cafeteria noise (+7 dB SNR). Audibility was calculated using linear and WDRC versions of the AAI. Subjects were required to adjust the volume control of their instrument until maximum clarity (intelligibility) was determined. Once the listener had indicated that a preferred volume-control setting had been achieved, the real-ear hearing aid output was measured using a probe-microphone system.

Results demonstrated no significant difference in speech audibility between amplification strategies for any speech level or between conditions in which competing noise was present or absent. Souza and Kitch (2001) also found that subjects tended to increase gain, as opposed to decreasing it, in the presence of background noise. The most important finding in this study, however, was that a large adjustment was needed in frequency-gain response, relative to the prescriptive formula, as the presentation level changed in intensity.

F. Prediction of Hearing Aid Outcomes—Non-Speech Benefits

Clinicians and researchers seem to agree that hearing loss assessment can be managed more effectively by using estimates of hearing handicap and benefit (McCarthy, 1994, 1998). As a result, there has been recent clinical interest in the relationship between AI and outcome measures. This interest stems from the ability of each procedure to quantify different aspects of auditory function. Outcome measures such as hearing-handicap scales, or inventories, differ from audiometric measures in that they provide information regarding the communication, social, emotional, and vocational consequences of the hearing impairment.

Holcomb *et al.* (2000) examined retrospectively the extent to which age, gender, degree of hearing sensitivity, and audiometric slope influence the relationship between perceived hearing handicap and AI. The files of 373 clinical patients between 18 and 85 years of age and those with normal hearing or acquired sensorineural hearing loss not exceeding 100 dB HL for the octave frequencies between 250 and 4000 Hz were included in the analyses. Audiometric slope was defined as the difference, in dB, between the best threshold for either ear at 4000 Hz minus the best threshold for either ear at 1000 Hz. All patient files included hearing-handicap scores derived by the Self-Assessment of Communication (SAC) (Schow and Nerbonne, 1982), and 292 of those also included scores from the Significant Other Assessment of Communication (SOAC) (Schow and Nerbonne, 1982). Although it is not evident from the article, it is assumed that these outcome measures were unaided. AI values were calculated using Humes' (1991) count-the-dot procedure and based on the best pure-tone data, regardless of ear, for octaves between 250 and 4000 Hz.

To determine the statistical relationships between AI and hearing-handicap scores, Pearson product-moment correlations were used. Additionally, partial correlations were also derived to determine the relationship between the AI, SAC, and SOAC for age, gender, hearing loss, and audiometric slope for individual and combined variables. Lastly, Spearman rank correlations were used to examine the relationship between each individual item on the SAC and SOAC and the AI.

Results revealed a significant negative Pearson product-moment correlation between AI and both measures of hearing handicap. This finding supports the premise that when speech information becomes less audible, hearing handicap increases. Furthermore, a statistically significant correlation was found between hearing-handicap measures. Partial-correlation analyses indicated that degree of hearing loss was the only variable found to influence the relationship between AI and hearing handicap. Analyses of AI values and individual items of each inventory demonstrated a significantly negative relationship, with communication-related items (items 1–6) correlating better with the AI than social-emotional items (items 7–10). These findings suggest that the AI and hearing-handicap measures each provides a unique view of a patient's communicative ability and thus should not be considered synonymous with respect to patient management.

Outcome measures have also been used to determine whether or not a hearing aid is providing listener benefit. Several investigators have suggested that the hearing aid that produces the highest aided AI score offers the best treatment choice for that patient (Pavlovic, 1989; Mueller and Killion, 1990; Humes, 1991). Because the best treatment is often dictated by measures made in environments not typical of everyday listening situations, other means are needed to augment what might be best for the patient. Outcome measures are one such tool. According to Souza and colleagues (2000), outcome measures can be described as measures of either efficacy or effectiveness. They defined efficacy as the degree of benefit that a specific group of patients experience from treatment under ideal conditions. Effectiveness, on the other hand, is the amount of benefit the average patient receives under real-world conditions.

Souza and colleagues (2000) determined recently the direct relationship between improved audibility and the overall effectiveness of hearing aids. They measured the effectiveness of global satisfaction and hearing aid adherence using two hearing-specific surveys and self-reported ratings. Subjects consisted of 115 patients seen through the Department of Veterans Affairs. All subjects were fit binaurally with custom analog, nonprogrammable instruments incorporating either peak-clipping or compression-limiting schemes. Target, unaided, and aided AI values were determined during the hearing aid fitting by means of real-ear analyzer software. The AI value was based on the $A_0(4)$ procedure developed by Pavlovic (1988) and modified to incorporate conversational speech levels (Pavlovic, 1991). The effectiveness of hearing aids was determined by the amount of change in listening ability without and with hearing aids. The Abbreviated Profile of Hearing Aid Benefit (APHAB) (Cox and Alexander, 1995) and the Satisfaction with Amplification in Daily Life (SADL) (Cox and Alexander, 1999) were the two standardized tests used. APHAB data were obtained unaided, prior to hearing aid fitting, and aided at the 1-month follow-up visit. Data for the SADL were obtained 1 year after fitting as a means to evaluate hearing aid quality. Adherence data were also gathered 1 year after fitting through the use of questions regarding the number of hours per day and days per week the aids were worn. Results revealed no systematic relationship between measures of improved audibility and patient ratings on communication ability. The authors hypothesized that this finding may be the result of the volume-control level at which measurements were initially made. A similar conclusion was also drawn for improved audibility and overall hearing aid satisfaction. A major outcome, however, was that a moderate relationship was found between hearing aid use and achievement of improved audibility.

G. Articulation Index-Directivity Index (AI-DI)

The AI can also be applied to directional-microphone hearing aids (DMHAs). Studies have shown that DMHAs effectively provide listeners with increased speech intelligibility in noise over omnidirectional-microphone devices (for a review, see Valente, 1999 and Amlani, 2001). This improvement in speech intelligibility occurs as a result of the DMHA's ability to attenuate sounds from the rear and sides of the listener with respect to sounds originating from directly in front. The amount of attenuation, however, can vary relative to microphone configuration and the aid's quantitatively determined directionality characteristics (Preves, 1997; Ricketts and Mueller, 1999; Valente, 1999, 2000; Ricketts and Dittberner, 2002). An electroacoustic method that can be used to account for these variations is the directivity index (DI). The DI is typically derived by presenting frequency-specific stimuli (often 500 to 4000 Hz) directly in front of the hearing aid as the aid is rotated either in free field or on a manikin at discrete angles, or azimuths. This electroacoustic measure is found in many manufacturers' technical specification manuals.

To provide a reasonable estimate of the improvement in speech recognition in noise for DMHAs under laboratory conditions, Killion and colleagues (1998) report a means by which audibility can be estimated from DI values. Using speech weighting from the Mueller and Killion (1990) count-the-dot audiogram, importance functions of 0.20, 0.23, 0.33, and 0.24 are assigned to the frequencies of 500, 1000, 2000, and 4000 Hz, respectively, to achieve audibility. This modification has been termed the Articulation Index-Directivity Index (AI-DI) and is calculated using the equation

$$\begin{aligned} AI\text{-}DI = {} & (0.2 \times DI_{500}) + (0.23 \times DI_{1000}) \\ & + (0.33 \times DI_{2000}) + (0.24 \times DI_{4000}) \end{aligned} \qquad (20)$$

The *AI-DI* is calculated first by determining each manufacturer-reported frequency-specific DI value. For example, assume a manufacturer reports values of 2.1, 2.5, 3, and 2.5, measured at 500, 1000, 2000, and 4000 Hz, respectively, for an in-the-ear instrument. Second, multiply each DI value by its frequency-specific importance function. Lastly, sum the importance functions across frequencies, resulting in an AI-DI value of 2.59 (0.42 + 0.58 + 0.99 + 0.60).

In contrast, an *unweighted* DI value is determined by adding the manufacturer-reported DI values and dividing by 4, or the number of frequencies. Using the same values in the earlier example, the calculation results in an unweighted DI value of 2.53 ([2.1 + 2.5 + 3 + 2.5]/4). Note that the AI-DI provides a benefit of 0.06 dB (2.59-2.53) over the unweighted DI method as a result of the greater importance given to the high frequencies, predominately the 2000 Hz region. Despite its intuitive and clinical appeal for comparing the improvement in speech recognition in noise for DMHAs, the AI-DI has not been investigated systematically.

H. Summary

Based on the evidence provided in this section, it can be concluded that:

1. No single prescriptive formula, either linear or nonlinear, has been determined to best maximize speech intelligibility. To date, recently developed prescriptive formulas have not been investigated empirically.
2. Almost any change to a hearing aid's frequency response results in a change in the AI value. Changes in the input level or hearing aid gain will also result in changes in the AI. Various electroacoustic features of a hearing aid interact with each other, making a determination of which electroacoustic features optimize the AI very difficult. Determining which electroacoustic features maximize speech-intelligibility performance, therefore, is an issue that will not be easily resolved.
3. Use of the AI as an outcomes-assessment tool seems enticing, provided the AI calculation is performed at the volume-control setting used in everyday life. We believe that further empirical evidence is needed, however, before clinical implementation can be validated.

Conclusion

Devised as a tool for predicting the intelligibility of speech transmitted by telecommunication devices, the Articulation Index and its successor, the Speech Intelligibility Index, have been refined in ways that make them clinically useful to audiologists. Using these indices, audiologists can theoretically predict unaided and aided speech intelligibility in their hearing-impaired patients without having to perform time-consuming measurements of intelligibility. It is mainly through count-the-dot audiograms that these predictive methods have found their way into the routine practice of clinical audiology. Implementation of these procedures as a primary tool in the fitting of hearing aids and in the management of hearing loss, however, has been limited. To date, the NAL-NL1 fitting method (Dillon, 1999; Byrne *et al.*, 2001), is the only formal prescriptive procedure developed specifically to use the principles of articulation theory to prescribe the hearing aid frequency-response characteristics that maximize speech intelligibility.

Great strides have been made to incorporate the AI in assessments of unaided and aided speech audibility for hearing-impaired listeners and to overcome inaccuracies inherent in these indices by attempting to account for a variety of confounding factors, such as severity of hearing loss, desensitization, age, distortion, and upward spread of masking. This review has made it obvious that these applications, to be useful for predicting intelligibility in hearing-impaired listeners, will require the incorporation of such corrections because they account for much of the variability in the speech intelligibility of hearing-impaired listeners with moderately severe-to-profound hearing impairments.

Researchers, to achieve their desired research objectives, have promulgated piecemeal formulas that have received rather limited use beyond the scope of their particular studies. For clinical use, it is desirable that corrections to the AI be combined, insofar as possible, into a single formula, and that such formulas be implemented through software applications that audiologists can readily apply. Clinically useful applications based on such refinements are likely to become a reality only if researchers, clinicians, and manufacturers of hearing aids, hearing aid analyzers, and real-ear analyzers collaborate to contribute their unique perspectives.

As Licklider and Miller (1951) noted—in the quote in our introduction—it is ostensibly more efficient to predict speech intelligibility than to measure it directly. Our review of the literature leads to the observation that research efforts to develop suitable models for *predicting* speech intelligibility also require substantial time and resources. As long as such efforts continue to show promise for utilization in unaided and aided applications for the hearing impaired, investigations of more efficient methods to determine speech audibility and to predict intelligibility appear justified.

Except for the various count-the-dot audiograms and the NAL-NL1 fitting approach, current measurement strategies remain relatively inefficient and difficult to implement clinically in demonstrating, either before or after a fitting, that a particular hearing aid or set of electroacoustic characteristics has potential for optimizing speech intelligibility for a given hearing-impaired listener. Because contemporary fitting strategies essentially allow the direct measurement of intelligibility only after the fitting process is completed, the need will continue for clinical procedures that predict or estimate speech-intelligibility outcomes for alternative hearing aids or signal-processing schemes. With further validation, the AI could also be used following the fitting of hearing aids as an outcome-assessment tool.

More research is needed to develop the traditional (auditory-only) AI before it can be used with confidence to predict speech-intelligibility performance under real-life conditions in individual hearing-impaired listeners. Investigation of the integration of auditory and speech-reading cues appears to be a natural extension of such research. Furthermore, the principles and procedures of the AI may prove to be extremely beneficial in the development of environmentally adaptive signal-processing schemes, as well as in future prescriptive formulas that optimize aided speech intelligibility.

Successful implementation of the AI in the audiology clinic depends not only on success in the laboratory but also upon the ability to build bridges between scientific theory and clinical applications. Count-the-dot audiograms are a good example of success in bridging the laboratory and the clinic. As is true of the past, such efforts in the future will likely evolve through the work of a core of interested researchers who tend to function somewhat independently. Ideally, a more systematic research approach is needed in which well-funded, collaborative research is aimed at developing relatively simple techniques that audiologists can use confidently in hearing aid selection and fitting. We would encourage the development and refinement of simplified procedures such as that described by Dillon (1993).

Investigations should, of course, aim at discovery of all those factors that affect the accuracy of intelligibility prediction, including those covered in this review (ie, hearing thresholds, speech and noise spectra, reverberation, desensitization, distortion, masking, visual cues). Indeed, if all or most of these factors can be addressed adequately in the predictive scheme, it would be reasonable to base predictions of intelligibility on either sentences or connected discourse only, as opposed to attempting to account for all other types of speech stimuli (eg, syllables and words). In addition, we view the development of appropriate psychoacoustic measures to augment audiometric thresholds as essential to improving the prospects of developing a clinical tool that can accurately predict speech intelligibility in individual hearing-impaired listeners. It is hoped that these attempts to produce an accurate predictor of intelligibility will be accompanied by parallel attempts to develop short—and scientifically valid—direct tests of speech intelligibility. We believe, though, that the many potential uses of an accurate predictor of intelligibility would ultimately make it a more efficient tool for widespread clinical use than any given direct test of intelligibility.

In the near future, it is hoped that research efforts will provide the kinds of information that will allow practical applications of AI measures to take a more prominent role in the practice of clinical audiology. These applications can be expected to improve significantly the diagnostic and rehabilitation capabilities of audiologists, allowing them to serve their hearing-impaired patients more effectively.

Acknowledgments

The authors gratefully acknowledge the contributions of Harvey Dillon, Todd Ricketts, and two anonymous reviewers for their comments on preliminary versions of the manuscript. We also express gratitude to our colleagues who unselfishly contributed their original figures.

Glossary

Articulation Index (AI)—The quantitative aspect of articulation theory that accounts for the contribution of audible speech cues in given frequency bands to speech intelligibility. Refers specifically to the principles and procedures described in the ANSI S3.5-1969 standard. Also known as Audibility Index.

Articulation Index-Directivity Index (AI-DI)—A means by which audibility can be estimated from directivity index (DI) values using speech weighting from the Mueller and Killion

(1990) count-the-dot audiogram. The AI-DI is believed to provide a reasonable estimate of the improvement in speech recognition in noise for directional-microphone hearing aids under laboratory conditions.

Articulation Theory—A model that assumes the intelligibility of speech can be described through any communication system using weighted measurements of the speech-frequency regions audible to the listener.

Audibility—The ability to detect the presence of a sound.

Audibility Function (A_i)—A variable that refers to the amount of speech energy that is above the listener's threshold and any competing noise in a given frequency band, based on a 30-dB dynamic range of speech.

Audibility Index (AI)—A generic term referring to the quantative aspect of articulation theory that acounts for the contribution of audible speech cues in given frequency bands to speech intelligibility. It encompasses both the Articulation Index and the Speech Intelligibility Index.

Count-the-Dot Audiogram—A simplified clinical approach to the Audibility Index that depicts frequency-importance functions (I_i) as a vertical series of dots superimposed on a conventional audiogram within a frequency-intensity area highlighted to indicate the spectral weighting of speech cues.

Desensitization—Reduction in the ability of an ear with severe-to-profound sensorineural hearing impairment to extract audible information that contributes to speech intelligibility. As a means to account for this reduction, correction factors may be applied to the pertinent standards (ANSI S3.5-1969 and ANSI S3.5-1997).

Frequency-Importance Function (I_i)—A stimulus-dependent variable that represents the relative contribution of different frequency bands to speech intelligibility. Contributions of these different frequency bands must equal a value of 1.0.

Modified Speech Transmission Index (mSTI)—A hybrid model that combines the STI's approach of determining signal-to-noise ratio with the AI's approach of using one-third octave bands. Uses weighting factors determined by French and Steinberg (1947).

Proficiency Factor (*P*)—A correction factor that accounts for variables relating to practice and experience of the talker and the listener. It is often used as a generalized variable for individual differences that are not explained by audibility.

Rapid Speech Transmission Index (RASTI)—A manufacturer-based derivation of the STI used to predict speech intelligibility in various rooms and other enclosed environments such as auditoriums, theaters, and concert halls.

Speech Intelligibility Index (SII)—The quantitative aspect of articulation theory that accounts for the contribution of audible speech cues in given frequency bands to speech intelligibility. Refers specifically to the principles and procedures described in the ANSI S3.5-1997 standard. Also known as Audibility Index.

Speech Recognition Sensitivity (SRS)—An intelligibility theory based on a macroscopic model incorporating statistical-decision theory, which predicts speech-intelligibility performance based on the redundant and synergetic interactions among the spectral components of speech.

Speech Transmission Index (STI)—An acoustical index that shares some of the same features as the AI. The STI differs in that it was developed as a model for predicting audibility under temporally distorted conditions (eg, reverberation) often found in the real world.

Appendix A
Calculation of Self-Speech Masking and Upward Spread of Masking

Self-speech masking, in which the internal energy within each speech band is masked, is represented as

$$V_i = E'_i - 24 \quad \text{(A1)}$$

where V_i is the spectrum level for self-speech masking in the ith band and E'_I represents the equivalent speech-spectrum level for the ith band. Spectrum level is defined, for a specified signal at a particular frequency, as the sound pressure level (SPL) of that part of a signal contained with a band 1 cycle wide and centered at that frequency. The equivalent speech-spectrum level is defined as the measured level of speech determined at the point corresponding to the center of the listener's head, with the listener absent, while producing at the ear drum of the listener the same sound pressure level that exists under actual listening conditions.[8]

The spread of masking across speech bands is also defined by a set of equations that differ based on the width of the band. When using octave bands, the formula used is

$$C_i = -80 + 0.6\ [B_i + 10\log\ (h_i - l_i)] \quad \text{(A2)}$$

In this formula, C_i characterizes the slope per octave of the spread of masking, h_i the upper limiting frequency of a given band, and l_i the lower limiting frequency of the same band. B_i is the larger of the spectrum levels between the equivalent noise spectrum level (N'), calculated similarly to E', but for the noise spectrum and self-speech masking (V).

For calculations based on one-third octave bands, the formula is nearly identical:

$$C_i = -80 + 0.6\ [B_i\ 10\log\ (F_i - 6.353)] \quad \text{(A3)}$$

Here, note that the variable F, or the center frequency of the band, has replaced the upper and lower band limits. As a result, using this method requires an 18-band method.

Appendix B
Calculation of Level Distortion Factor

The level distortion factor (L_i) accounts for the decrease in speech intelligibility at high presentation levels. L_I is expressed as:

$$L_I = 1 - (E'_i - U_i - 10)/160 \quad \text{(B1)}$$

where E'_i is the equivalent speech-spectrum level and U_i is the standard speech-spectrum level at normal vocal effort. The equivalent speech-spectrum level is defined as the measured level of speech determined at the point corresponding to the center of the listener's head, with the listener absent, while producing at the ear drum of the listener the same sound pressure level that exists under actual listening conditions (see footnote in Appendix A for more details). The standard speech spectrum is a measurement of speech made directly in front of the talker's lips in quiet and at a distance of 1 meter in a sound field. U_i also assumes normal vocal effort, which is determined at an overall sound pressure level of 62.35 dB, but may also account for raised, loud, and shouted vocal efforts.

[8]An actual listening situation, as described in the ANSI S3.5-1997 standard, refers to either monaural listening or the same signal reaching the left and right ears. For binaural conditions that fall outside this definition, the standard incorporates a correction factor.

Appendix C
An Example of the Calculation Method for the STI

	Octave-Band Center Frequency (Hz)						
	125	250	500	1000	2000	4000	8000
MTF (Converted to SNR)	-4.5	4.5	11.2	15.0	15.0	15.0	15.0
	-2.6	3.6	12.0	15.0	15.0	15.0	15.0
	-5.3	3.5	11.6	14.6	15.0	15.0	15.0
	-1.5	2.8	9.9	15.0	15.0	15.0	15.0
	-2.7	5.3	10.0	15.0	15.0	15.0	15.0
	-3.1	1.9	12.1	14.1	15.0	15.0	15.0
	-3.8	2.5	11.9	15.0	15.0	15.0	15.0
	-4.0	5.0	10.6	15.0	15.0	15.0	15.0
	-3.6	3.8	12.3	13.8	15.0	15.0	15.0
	-4.2	4.3	13.1	15.0	15.0	15.0	15.0
	-2.3	3.9	12.3	14.3	15.0	15.0	15.0
	-1.9	4.1	11.5	14.5	15.0	15.0	15.0
	-2.9	3.8	12.0	15.0	15.0	15.0	15.0
	-3.3	5.0	13.8	13.8	15.0	15.0	15.0
	-2.8	4.8	11.7	15.0	15.0	15.0	15.0
Σ MTF	-48.5	57.8	176.0	220.1	225.0	225.0	225.0
Average MTF	-3.2	3.9	11.7	14.7	15.0	15.0	15.0
Average MTF Re: Dynamic Range of Speech	11.8	18.9	26.7	29.3	30.0	30.0	30.0
Transmission Index	0.39	0.63	0.89	0.98	1.0	1.0	1.0
Weighting Factor	0.129	0.143	0.114	0.114	0.186	0.171	0.143
Octave-Weighted Values	0.05	0.09	0.10	0.11	0.19	0.17	0.14
			STI =0.85				

The Speech Transmission Index (STI) can be determined either by direct measurement or through a calculation method. Only the calculation method is described here. Readers interested in details regarding the direct-measurement procedure are referred to the Steeneken and Houtgast (1980) article.

To calculate the STI, modulation transfer functions (MTFs) are determined for the octave frequencies 125 Hz to 8000 Hz at each of the following modulation frequencies: 0.63, 0.80, 1.00, 1.25, 1.60, 2.00, 2.50, 3.15, 4.00, 5.00, 6.30, 8.00, 10.00, and 12.50 Hz. These MTFs are then converted to signal-to-noise ratios (SNRs), as described in the text and shown in the table. In the example presented, these values are hypothetically derived. Summing the values and dividing by the denominator (ie, the number of MTFs) will result in the average MTF for a given octave. Next, the average MTFs per octave are used to calculate the dynamic range of speech available. This is achieved by adding 15 at each octave's average MTF. The average MTFs for the nominal dynamic range of speech is then divided by 30, or the absolute dynamic range of speech. This yields a proportion, or the Transmission Index per octave. To calculate the predictive transmission of speech within this acoustical environment, the Transmission Index at each octave is multiplied by a weighting factor shown to be optimal for correlating the STI to monosyllabic-word scores. Notice that the weighting factors across frequency equal to 1 and are essentially equal across all octave bands. The last step in the STI calculation method requires the summing of the octave-weighted values. In this case, the STI resulted in a value of 0.85.

References

Abel SM, Sass-Kortsak A, Naugler JJ. The role of high-frequency hearing in age-related speech understanding deficits. Scand Audiol 29:131-138, 2000.

Acoustical Society of America. American standard for noise measurement (Z24.2-1942). J Acoust Soc Am 13:102, 1942.

Allen JB, Hall JL, Jeng PS. Loudness growth in ½ octave bands (LGOB)–A procedure for the assessment of loudness. J Acoust Soc Am 88:745-753, 1990.

American National Standards Institute. ANSI S3.5-1969. Methods for the Calculation of the Articulation Index. New York: Acoustical Society of America, 1969.

American National Standards Institute. ANSI S3.14-1977. Rating Noise with Respect to Speech Interference. New York: Acoustical Society of America, 1977.

American National Standards Institute. ANSI S3.21-1978 (R1997). Method for Manual Pure-Tone Threshold Audiometry. New York: Acoustical Society of America, 1978.

American National Standards Institute. ANSI S3.2-1989. Methods for Measuring the Intelligibility of Speech Over Communication Systems. New York: Acoustical Society of America, 1989.

American National Standards Institute. ANSI S3.6-1989. Specifications for Audiometers. New York: Acoustical Society of America, 1989.

American National Standards Institute. ANSI S3.5 Draft v.3.1-1993. Methods for the Calculation of the Articulation Index. New York: Acoustical Society of America, 1993.

American National Standards Institute. ANSI S1.1-1994 (R1999). Acoustical Terminology. New York: Acoustical Society of America, 1994.

American National Standards Institute. ANSI S3.5-1997. Methods for the Calculation of the Speech Intelligibility Index. New York: Acoustical Society of America, 1997.

Amlani AM. Directional microphone hearing aids: A meta-analytic perspective. J Am Acad Audiol 12:202-214, 2001.

Bacon SP, Gleitman RM. Modulation detection in subjects with relatively flat hearing losses. J Speech Hear Res 35:642-653, 1992.

Bacon SP, Viemeister NF. Temporal modulation function in normal-hearing and hearing-impaired listeners. Audiology 24:115-134, 1985.

Bell ST, Dirks DD, Trine TD. Frequency-importance functions for words in high- and low-context sentences. J Speech Hear Res 35:950-959, 1992.

Bench RJ, Doyle JM. The BKB/A (Bamford-Kowal-Bench/Australian Version) Sentence Lists for Hearing-Impaired Children. La Trobe University, Victoria, 1979.

Bentler RA, Pavlovic CV. Transfer functions and correction factors used in hearing aid evaluation and research. Ear Hear 10:58-63, 1989.

Beranek LL. The design of speech communication systems. Proc Inst Radio Eng 35:880-890, 1947.

Bilger RC. Speech recognition test development. Asha Report 14, 1985.

Binnie CA. Bi-sensory articulation factors for normal hearing and sensorineural hearing loss patients. J Acad Rehab Audiol 3:43-52, 1973.

Binnie CA, Montgomery AA, Jackson PL. Auditory and visual contributions to the perception of consonants. J Speech Hear Res 17:619-630, 1974.

Black JW. Equally contributing frequency bands in intelligibility testing. J Speech Hear Res 2:81-83, 1959.

Boothroyd A. Developments in speech audiometry. Sound 2:3-10, 1968.

Braida LD. Discussion summary V (Chapters 14–17): Hearing aid response characteristic selection strategies. In: Studebaker GA, Hochberg I, eds: Acoustical Factors Affecting Hearing Aid Performance. Baltimore: University Park Press, 1980, pp. 424-430.

Braida LD, Durlach NI, Lippman RP, Hicks BL, Rabinowitz WM, Reed CM. Hearing aids: A review of past research on linear amplification, amplitude compression and frequency lowering. Asha Monograph Number 19, 1979.

Brüel & Kjaer. The modulation transfer function in room acoustics. RASTI: A tool for evaluating auditoria. Technical Review No 3-1985. MA: Brüel & Kjaer, 1985.

Burkhard MD, Sachs RM. Anthropometric manikin for acoustic research. J Acoust Soc Am 58:214-222, 1975.

Byrne D. Effects of bandwidth and stimulus type on most comfortable loudness levels of hearing-impaired listeners. J Acoust Soc Am 80:484-493, 1986a.

Byrne D. Effects of frequency response characteristics on speech discrimination and perceived intelligibility and pleasantness of speech for hearing-impaired listeners. J Acoust Soc Am 80:494-504, 1986b.

Byrne D. Key issues in hearing aid selection and evaluation. J Am Acad Audiol 3:67-80, 1992.

Byrne D, Cotton S. Preferred listening levels of persons with sensorineural hearing losses. Aust J Audiol 9:7-14, 1987.

Byrne D, Dillon H. The National Acoustic Laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear 7:257-265, 1986.

Byrne D, Parkinson A, Newall P. Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired ear. Ear Hear 11:27-36, 1990.

Byrne D, Dillon H, Ching T, Katsch R, Keidser G. NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. J Am Acad Audiol 12:37-51, 2001.

Cavanaugh WJ, Farrell WR, Hirtle PW, Walters GB. Speech privacy in buildings. J Acoust Soc Am 34:475-492, 1962.

Ching T, Dillon H, Byrne D. Prediction of speech recognition from audibility and psychoacoustic abilities of hearing impaired listeners. In: Jesteadt W, ed: Modeling Sensorineural Hearing Loss. Mahwah, New Jersey: Lawrence Erlbaum, 1997, pp. 433-446.

Ching TY, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. J Acoust Soc Am 103:1128-1140, 1998.

Ching TYC, Dillon H, Katsch R, Byrne D. Maximizing effective audibility in hearing aid fitting. Ear Hear 22:212-224, 2001.

Cornelisse LE, Seewald RC, Jamieson DG. Wide dynamic range compression hearing aids: The DSL[i/o] approach. Hear J 47(10):23-29, 1994.

Cornelisse LE, Seewald RC, Jamieson DG. The input/output formula: A theoretical approach to the fitting of personal amplification devices. J Acoust Soc Am 97:1854-1864, 1995.

Cox R, Alexander G. The abbreviated profile of hearing aid benefit. Ear Hear 16:176-183, 1995.

Cox RM, Alexander GC. Measuring satisfaction with amplification in daily life: The SADL scale. Ear Hear 20:306-320, 1999.

Cox RM, McDaniel DM. Development of the speech intelligibility rating (SIR) test for hearing aid comparisons. J Speech Hear Res 32:347-352, 1989.

Crain TR, Van Tasell D. The effect of peak clipping on the speech recognition threshold. Ear Hear 15:443-453, 1994.

DePaolis RA, Janota CP, Frank T. Frequency importance functions for words, sentences, and continuous discourse. J Speech Hear Res 39:714-723, 1996.

Dillon H. Hearing aid evaluation: Predicting speech gain from insertion gain. J Speech Hear Res 36:621-633, 1993.

Dillon H. NAL-NL1: A new procedure for fitting non-linear hearing aids. Hear J 52(4):10-16, 1999.

Dubno JR, Dirks DD, Ellison DE. Stop-consonant recognition for normal-hearing listeners with high-frequency hearing loss: I. The contributions of selected frequency regions. J Acoust Soc Am 85:347-354, 1989a.

Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am 14:275-284, 1984.

Dubno JR, Dirks DD, Schaefer AB. Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II: Articulation index predictions. J Acoust Soc Am 85:355-364, 1989b.

Dugal RL, Braida, LD, Durlach NI. Implications of previous research for the selection of frequency-gain characteristics. In: Studebaker GA, Hochberg I, eds: Acoustical Factors Affecting Hearing Aid Performance. Baltimore: University Park Press, 1980, pp. 379-403.

Duggirala V, Studebaker GA, Pavlovic CV, Sherbecoe RL. Frequency importance functions for a feature recognition test material. J Acoust Soc Am 83:2372-2382, 1988.

Dunn HK, White SD. Statistical measurements on conversational speech. J Acoust Soc Am 11:278-288, 1940.

Duquesnoy AJ, Plomp R. Effect of reverberation and noise on the intelligibility of sentences in cases of presbycusis. J Acoust Soc Am 68:537-544, 1980.

Egan JP. Articulation testing methods. Laryngoscope 58:955-991, 1948.

Erber NP. Interaction of audition and vision in the recognition of oral speech stimuli. J Speech Hear Res 12:423-425, 1969.

Fabry DA, Van Tasell DJ. Evaluation of an articulation-index based model for predicting the effects of adaptive frequency response hearing aids. J Speech Hear Res 33:676-689, 1990.

Festen JM, Van Dijkhuizen JN, Plomp P. Considerations on adaptive gain and frequency response in hearing aids. Acta Otolaryngol 469(Suppl) :193-201, 1990.

Finitzo-Heiber T, Tillman TW. Room acoustic effects on monosyllabic word discrimination ability for normal and hearing-impaired children. J Speech Hear Res 21:440-458, 1978.

Fletcher H. The perception of speech sounds by deafened persons. J Acoust Soc Am 24:490-497, 1952.

Fletcher H. Speech and Hearing in Communication. New York: Krieger, 1953.

Fletcher H, Galt RH. The perception of speech and its relation to telephony. J Acoust Soc Am 22:89-151, 1950.

Fletcher H, Munson WA. Masking and critical bands. J Acoust Soc Am 9:1, 1937.

Fletcher H, Steinberg JC. Articulation testing methods. Bell Sys Tech J 8:806-854, 1929.

French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. J Acoust Soc Am 19:90-119, 1947.

Gates GA, Popelka GR. Neural presbycusis: A diagnostic dilemma. Am J Otol 13:313-317, 1992.

Grant KW, Braida LD. Evaluating the articulation index for auditory-visual input. J Acoust Soc Am 89:2952-2960, 1991.

Green DM, Birdsall TG. The effect of vocabulary size on articulation score. Technical Memorandum No. 81

and Technical Note AFCRC-TR-57-58, University of Michigan: Electronic Defense Group, 1958.

Hargus SE, Gordon-Salant S. Accuracy of speech intelligibility index predictions for noise-masked young listeners with normal hearing and for elderly listeners with hearing impairment. J Speech Hear Res 38:234-243, 1995.

Hawkins DB, Montgomery AA, Mueller HG, Sedge RK. Assessment of speech intelligibility by hearing-impaired listeners. In: Berglund B, Berglund U, Karlson J, Lindwall T, eds: Noise as a Public Health Problem. Stockholm: Swedish Council for Building Research, 1988, pp. 241-246.

Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. J Speech Hear Disord 17:321-337, 1952.

Hogan CA, Turner CW. High-frequency audibility: Benefits for hearing-impaired listeners. J Acoust Soc Am 104:432-441, 1998.

Hohmann V, Kollmeier B. The effect of multichannel dynamic compression on speech intelligibility. J Acoust Soc Am 97:1191-1195, 1995.

Holcomb LM, Nerbonne MA, Konkle DF. Articulation index and hearing handicap. J Am Acad Audiol 11:224-229, 2000.

Hou Z, Thornton AR. A model to evaluate and maximize hearing aid performance by integrating the articulation index across listening conditions. Ear Hear 15:106-112, 1994.

Houtgast T, Steeneken HJM. Evaluation of speech transmission channels by using artificial signals. Acustica 25:355-367, 1971.

Houtgast T, Steeneken HJM. The modulation transfer function in room acoustics as a predictor of speech intelligibility. Acustica 28:66-73, 1973.

Houtgast T, Steeneken HJM. Experimental verification of the STI: A valid basis for setting indoor noise level criteria. In Rossi G, ed: Proceedings of the Fourth International Congress on Noise as a Public Health Problem. Turin, Italy: CSRA, 1983, pp. 477-487.

Houtgast T, Steeneken HJM. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am 77:1069-1077, 1985.

Houtgast T, Steeneken HJM, Plomp R. Predicting speech intelligibility in rooms from the modulation transfer function: I. General room acoustics. Acustica 46:60-72, 1980.

Humes LE. An evaluation of several rationales for selecting hearing aid gain. J Speech Hear Disord 51:272-281, 1986.

Humes LE. Understanding the speech-understanding problems of the hearing impaired. J Am Acad Audiol 2:59-69, 1991.

Humes LE. Some temporal factors affecting speech recognition. In Studebaker GA, Hochberg I, eds: Acoustical Factors Affecting Hearing Aid Performance (2nd ed). Boston: Allyn and Bacon, 1993, pp. 205-220.

Humes LE, Coughlin M, Talley L. Evaluation of the use of a new compact disc for auditory perceptual assessment in the elderly. J Am Acad Audiol 7:419-427, 1996.

Humes LE, Dirks DD, Bell TS, Ahlstrom C, Kincaid GE. Application of the articulation index to the recognition of speech by normal-hearing and hearing-impaired listeners. J Speech Hear Res 29:447-462, 1986.

Humes LE, Halling DC. Overview, rationale, and comparison of suprathreshold-based gain prescription methods. In: Valente M, ed: Strategies for Selecting and Verifying Hearing Aid Fittings. New York: Thieme, 1994, pp. 19-37.

Humes LE, Riker S. Evaluation of two clinical versions of the articulation index. Ear Hear 13:406-409, 1992.

Humes LE, Roberts L. Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility. J Speech Hear Res 33:726-735, 1990.

Humes LE, Watson BU, Christensen LA, Cokely CG, Halling DC, Lee L. Factors associated with individual differences in clinical measures of speech recognition among the elderly. J Speech Hear Res 37:465-473, 1994.

Jerger J, Jerger S, Oliver T, Pirozzolo F. Speech understanding in the elderly. Ear Hear 10:79-89, 1989.

Kamm CA, Dirks DD, Bell TS. Speech recognition and the articulation index for normal and hearing-impaired listeners. J Acoust Soc Am 77:281-288, 1985.

Killion M. The SIN report: Circuits haven't solved the hearing-in-noise problem. Hear J 50(10):28-30, 32, 34, 1997.

Killion M, Christensen L. The case of the missing dots: AI and SNR. Hear J 51(5):32-47, 1998.

Killion M, Schulerin R, Christensen L, Fabry D, Revit L, Niquette P, Chung K. Real-world performance of an ITE directional microphone. Hear J 51(4):24-26, 30, 32, 35, 36, 38, 1998.

Killion MC. The noise problem: There's hope. Hear Inst 36(11):26, 28, 30, 32, 1985.

Killion MC. New thinking on hearing in noise: A generalized articulation index. Seminars in Hearing 23:57-75, 2002.

Killion MC, Fikret-Pasa S. The 3 types of sensorineural hearing loss: Loudness and intelligibility considerations. Hear J 46(11):31-36, 1993.

Killion MC, Mueller HG, Pavlovic CV, Humes LE. A is for audibility. Hear J 46(4):29-32, 1993.

Knudsen VO. The hearing of speech in auditoriums. J Acoust Soc Am 1:56-82, 1929.

Kringlebotn M. A graphical method for calculating the speech intelligibility index and measuring hearing disability from audiograms. Scand Audiol 28:151-160, 1999.

Kryter KD. Methods for the calculation of the articulation index. J Acoust Soc Am 34:1689-1697, 1962a.

Kryter KD. Validation of the articulation index. J Acoust Soc Am 34:1698-1702, 1962b.

Leavitt R, Flexor C. Speech degradation as measured by the rapid speech transmission index (RASTI). Ear Hear 12:115-118, 1991.

Leijon A, Eriksson-Mangold M, Beck-Karlsen A. Preferred hearing aid gain and bass-cut of speech by hearing-impaired listeners. J Acoust Soc Am 77:281-288, 1984.

Liberman MC, Dodds LW. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. Hear Res 16:55-74, 1984.

Licklider JCR, Miller GA. The perception of speech. In: Stevens SS, ed: Handbook of Experimental Psychology. New York: John Wiley & Sons, Inc, 1951, pp. 1040-1074.

Lundeen C. Count-the-dot audiogram in perspective. Am J Audiol 5:57-58, 1996.

Magnusson L. Predicting the speech recognition performance of elderly individuals with sensorineural impairment: A procedure based on the speech intelligibility index. Scan Audiol 25:215-222, 1996.

Magnusson L, Karlsson M, Leijon A. Predicted and measured speech recognition performance in noise with linear amplification. Ear Hear 22:46-57, 2001.

McCandless GA, Lyregarrd PE. Prescription of gain/output (POGO) for hearing aids. Hear Inst 35(1):16-21, 1983.

McCarthy PA. Self-assessment inventories: They're not just for aural rehab anymore. Hear J 45(6):41-43, 1994.

McCarthy PA. Self-assessment revisited. Hear J 51(3):10-18, 1998.

Moncur JP, Dirks D. Binaural and monaural speech intelligibility in reverberation. J Speech Hear Res 10:186-195, 1967.

Moore BCJ. Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. Ear Hear 17:133-161, 1996.

Moore BCJ. A test for the diagnosis of dead regions in the cochlea. Brit J Audiol 34:205-224, 2000.

Moore BCJ. Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids. Trends Ampl 5:1-34, 2001.

Moore BCJ, Alcántara JI. The use of psychophysical tuning curves to explore dead regions in the cochlea. Ear Hear 22:268-278, 2001.

Mueller HG. Insertion gain measurements. In: Mueller HG, Hawkins DB, Northern JL, eds: Probe Microphone Measurements: Hearing Aid Selection and Assessment. San Diego: Singular Publishing, 1992, pp. 113-144.

Mueller HG, Killion MC. An easy method for calculating the articulation index. Hear J 43(9):14-17, 1990.

Murray N, Byrne D. Performance of hearing-impaired and normal hearing listeners with various high-frequency cutoffs in hearing aids. Aust J Audiol 8:21-28, 1986.

Musch H, Buus S. Using statistical decision theory to predict speech intelligibility. I. Model structure. J Acoust Soc Am 109:2896-2909, 2001a.

Musch H, Buus S. Using statistical decision theory to predict speech intelligibility. II. Measurement and prediction of consonant-discrimination performance. J Acoust Soc Am 109:2910-2920, 2001b.

Nabelek AK, Letowski TR, Tucker FM. Reverberant overlap- and self-masking in consonant identification. J Acoust Soc Am 86:1259-1265, 1989.

Nabelek AK, Nabelek IV. Room acoustics and speech perception. In: Katz J, ed: Handbook of Clinical Audiology (4th Ed.). Baltimore: Williams & Wilkins, 1994, pp. 624-637.

Nabelek AK, Pickett JM. Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners. J Speech Hear Res 17:724-739, 1974.

Nabelek AK, Robinson PK. Monaural and binaural speech perception in reverberation for listeners of various ages. J Acoust Soc Am 71:1242-1248, 1982.

Neuman AC, Hochberg I. Children's perception of speech in reverberation. J Acoust Soc Am 73:2145-2149, 1983.

Pascoe DP. Clinical implications of nonverbal method of hearing aid selection and fitting. Seminars Speech Hear Res 1:217-229, 1980.

Pavlovic CV. Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment. J Acoust Soc Am 75:1253-1258, 1984.

Pavlovic CV. Derivation of primary parameters and procedures for use in speech intelligibility predictions. J Acoust Soc Am 82:413-422, 1987.

Pavlovic CV. Articulation index predictions of speech intelligibility in hearing aid selection. Asha 30 June–July:63-65, 1988.

Pavlovic CV. Speech spectrum considerations and speech intelligibility predictions in hearing aid evaluations. J Speech Hear Disord 54:3-8, 1989.

Pavlovic CV. Speech recognition and five articulation indexes. Hear Inst 42(9):20, 22, 23, 1991.

Pavlovic CV. Band importance functions for audiological applications. Ear Hear 15:100-104, 1994.

Pavlovic CV, Studebaker GA. Evaluation of some assumptions underlying the articulation index. J Acoust Soc Am 75:1606-1612, 1984.

Pavlovic CV, Studebaker GA, Sherbecoe RL. An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals. J Acoust Soc Am 80:50-57, 1986.

Payton KL, Uchanski RM, Braida LD. Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. J Acoust Soc Am 95:1581-1592, 1994.

Pearson KS, Bennett RL, Fidell S. Speech Levels in Various Environments (BBN Report No. 3281). Cambridge, MA: Bolt, Beranek and Newman, 1976.

Pekkarinen E, Viljanen V. Acoustic conditions for speech communication in classrooms. Scand Audiol 20:257-263, 1991.

Pluvinage V. Rationale and development of the Resound system. In: Sandlin RE, ed: Understanding Digitally Programmable Hearing Aids. Boston: Allyn and Bacon, 1994, pp. 15-39.

Popelka GR, Mason D. Factors which affect the measures of speech audibility with hearing aids. Ear Hear 8(5S):109-118, 1987.

Plomp R, Mimpen AM. Speech reception threshold for sentences as a function of age and noise level. J Acoust Soc Am 66:1333-1342, 1979.

Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. J Acoust Soc Am 83:2322-2327, 1988.

Preves D. Directional microphone use in ITE hearing instruments. Hear Rev 4(7):21, 22, 24-27, 1997.

Rankovic CM. An application of the articulation index to hearing aid fitting. J Speech Hear Res 34:391-402, 1991.

Resnick SB, Dubno JR, Hoffnung S, Levitt H. Phoneme errors on a nonsense syllable test. J Acoust Soc Am 58(Suppl):S114, 1975.

Ricketts T, Mueller HG. Making sense of directional microphone hearing aids. Am J Audiol 8:117-127, 1999.

Ricketts TA, Dittberner AB. Directional amplification for improved signal-to-noise ratio: Strategies, measurement, and limitations. In: Valente M, ed: Strategies for Selecting and Verifying Hearing Aid Fittings (2nd ed). New York: Thieme Medical Publishers, 2002, pp. 274-346.

Schow RL, Nerbonne MA. Communication screening profile: Use with elderly clients. Ear Hear 3:135-147, 1982.

Schum DJ, Matthews LJ, Lee F. Actual and predicted word recognition performance of elderly hearing-impaired listeners. J Speech Hear Res 34:636-642, 1991.

Skinner MW. Speech intelligibility in noise in noise-induced hearing loss: Effects of high frequency compensation. J Acoust Soc Am 67:306-317, 1980.

Skinner MW. Hearing Aid Evaluation. Englewood Cliffs, New Jersey: Prentice Hall, 1988.

Souza PE, Bishop RD. Improving audibility with nonlinear amplification for listeners with high-frequency loss. J Am Acad Audiol 11:214-223, 2000.

Souza PE, Kitch VJ. Effect of preferred volume setting on speech audibility in different hearing aid circuits. J Am Acad Audiol 12:415-422, 2001.

Souza PE, Turner CW. Multichannel compression, temporal cues and audibility. J Speech Hear Res 41:315-326, 1998.

Souza PE, Turner CW. Quantifying the contribution of audibility to recognition of compressed-amplified speech. Ear Hear 20:12-20, 1999.

Souza PE, Yueh B, Sarubbi M, Loovis CF. Fitting hearing aids with the articulation index: Impact on hearing aid effectiveness. J Rehabil R D 37:473-481, 2000.

Steele JA, Binnie CA, Cooper WA. Combining auditory and visual stimuli in the adaptive testing of speech discrimination. J Speech Hear Disord 43:115-122, 1978.

Steeneken HJM, Houtgast T. A physical method for measuring speech-transmission quality. J Acoust Soc Am 67:318-326, 1980.

Steeneken HJM, Houtgast T. Mutual dependence of octave-band weights in predicting speech intelligibility. Speech Commun 28:109-123, 1980.

Stelmachowicz P, Lewis D, Kalberer A, Creutz T. Situational Hearing-Aid Response Profile for Users (SHARP, v 2.0). Omaha, Nebraska: Boy's Town National Research Hospital, 1994.

Stelmachowicz PG, Dalzell S, Petersen D, Kopun J, Lewis DL, Hoover BE. A comparison of threshold-based fitting strategies for nonlinear hearing aids. Ear Hear 19:131-138, 1998.

Studebaker GA. A 'rationalized' arcsine transform. J Speech Hear Res 28:455-462, 1985.

Studebaker GA. The effect of equating loudness on audibility-based hearing aid selection procedures. J Am Acad Audiol 3:113-118, 1992.

Studebaker GA, McDaniel DM, Sherbecoe RL. Evaluating relative speech recognition performance using the proficiency factor and rationalized arcsine differences. J Am Acad Audiol 6:173-182, 1995.

Studebaker GA, Pavlovic CV, Sherbecoe RL. A frequency importance function for continuous discourse. J Acoust Soc Am 81:1130-1138, 1987.

Studebaker GA, Sherbecoe RL. Frequency-importance and transfer functions for recorded CID W-22 word lists. J Speech Hear Res 34:427-438, 1991.

Studebaker GA, Sherbecoe RL. Frequency-importance functions for speech recognition. In: Studebaker GA, Hochberg I, eds: Acoustical Factors Affecting Hearing Aid Performance (2nd ed). Boston: Allyn and Bacon, 1993, pp. 185-204.

Studebaker GA, Sherbecoe RL. Intensity-importance functions for bandlimited monosyllabic words. J Acoust Soc Am 111:1422-1436, 2002.

Studebaker GA, Sherbecoe RL, Gilmore C. Frequency-importance and transfer functions for the Auditec of Saint Louis recording of the NU-6 word test. J Speech Hear Res 36:799-807, 1993.

Studebaker GA, Sherbecoe RL. Evaluating the Speech Recognition Performance of Hearing-Impaired Subjects. Paper presented at the Lake Arrowhead International Conference on Issues on Advanced Hearing Aid Research, Lake Arrowhead, California, May 30–June 4, 1994.

Studebaker GA, Sherbecoe RL, McDaniel DM, Gwaltney CA. Monosyllabic word recognition at higher-than-normal speech and noise levels. J Acoust Soc Am 105:2431-2444, 1999.

Studebaker GA, Sherbecoe RL, McDaniel DM, Gray GA. Age-related changes in monosyllabic word recognition when audibility is held constant. J Am Acad Audiol 8:150-162, 1997.

Sumby WH, Pollack I. Visual contributions to speech intelligibility in noise. J Acoust Soc Am 26:212-215, 1954.

Tillman TW, Carhart R. An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words. Northwestern University Auditory Test No.6. (USAF School of Aerospace Medicine Technical Report). Brooks Air Force Base, Texas, 1966.

Turner CW, Cummings KJ. Speech audibility for listeners with high-frequency hearing loss. Am J Audiol 8:150-162, 1999.

Turner CW, Robb MP. Audibility and recognition of stop consonants in normal and hearing-impaired subjects. J Acoust Soc Am 81:1566-1573, 1987.

Valente M. Use of microphone technology to improve user performance in noise. Trends in Amplification 4:112-135, 1999.

Valente M. Use of microphone technology to improve user performance in noise. In: Sandlin RE, ed: Textbook of Hearing Aid Amplification: Technical and Clinical Considerations (2nd ed). San Diego: Singular Thomason Learning, 2000, pp. 247-284.

van Rooij JCGM, Plomp R. The effect of linguistic entropy on speech perception in noise in young and elderly listeners. J Acoust Soc Am 90:2985-2991, 1991.

Van Tasell DJ. Hearing loss, speech, and hearing aids. J Speech Hear Res 36:228-244, 1993.

Villchur E. Comments on "The negative effect of amplitude compression on multichannel hearing aids in the light of the modulation transfer function" (Plomp. J Acoust Soc Am 83:2322-2327, 1988). J Acoust Soc Am 86:425-427, 1989.

Waring CE. Transformation and Reception of Sounds Under Combat Conditions. Office of Scientific Research and Development #TIPU 1130. Washington DC: Office of Naval Research, 1946.

Webster J. Interpretations of speech and noise characteristics of NTID learning centers. J Acoust Soc Am 66(Suppl):S37, 1979.

Wilde G, Humes LE. Applications of the articulation index to the speech recognition of normal and impaired listeners wearing hearing protection. J Acoust Soc Am 87:1192-1199, 1990.

Willot JF. Aging and the Auditory System: Anatomy, Physiology, and Psychophysics. San Diego: Singular Publishing, Inc, 1991.

Zurek PM, Delhorne LA. Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. J Acoust Soc Am 82:1548-1559, 1987.